2006

Critical Tax Theory Conference

Macon, Georgia
April 7 & 8, 2006

Conference Materials

The 2006 Critical Tax Theory Conference is sponsored by generous funding from Daisy Floyd, Dean and Professor at Mercer University School of Law
# Table of Contents

A. General Conference Information  
   
B. Conference Program  
   
C. List of Conference Participants  
   
D. Biography of Keynote Speaker  
   
E. Biographies of Conference Participants  
   
F. Conference Papers  

1  
4  
8  
10  
11  
18
A. General Conference Information

The Conference: This year’s Critical Tax Theory Conference will take place at Mercer University School of Law in Macon, Georgia, on Friday and Saturday, April 7 and 8, 2006. We plan to start the conference promptly at 1:00 pm on Friday, April 7, 2006, and end promptly at 5:00 pm on Saturday, April 8, 2006. In addition, the law school will treat all conference participants to dinner on Friday night, lunch on Saturday afternoon and dinner again on Saturday night. All relevant travel and lodging information is outlined below.

We will be honored this year with a special address by Nina Olson, National Taxpayer Advocate for the Internal Revenue Service. Ms. Olson will address our group on Friday evening, April 7.

In addition to Ms. Olson’s address, we will have all traditional components of the conference, including formal presentations of works-in-progress (1 hour to 1.5 hour panels of 2-3 related projects per panel) and informal incubator sessions (5-15 minute round-table style presentations). The conference program, included in these materials, outlines the various panels and participants therein. Finally, if you are a panelists and have any needs for presentation equipment (document camera, power point, etc.), please let Professor David A. Brennen know immediately.

Location of Conference: The 2006 Critical Tax Theory Conference will be held at Mercer University School of Law in Macon, Georgia. The greater Macon area has a population of approximately 350,000. The historical city of Macon is located approximately 80 miles south of Atlanta. Founded in 1833, Mercer University is a private, co-educational university with an enduring Baptist heritage. The University’s impressive main administration building is listed on the National Register of Historic Places. The main law school building was constructed as a replica of Independence Hall. Mercer Law School is noted for having the nation’s #1 ranked legal writing program and for having received two awards in the past decade from the American Bar Association for its curriculum – both for the overall design and for the inclusion of a required course on professionalism.

Transportation to the Conference: The closest major airport to Macon is Atlanta Hartsfield-Jackson International Airport (ATL) – the busiest airport in the world. Transportation from the Atlanta airport to Macon is provided by Groome Transportation Service. Groome Transportation Service leaves the airport every hour on the hour (and sometimes on the half hour) en route to Macon and the Crown Plaza Hotel – the conference hotel. The cost is $27 one-way or $50 round trip. Groome accepts all major credit cards, except Discover. The toll free number is 800-537-7906. An alternate arrival airport is the Macon Regional Airport (MCN). Although it is often more expensive than the Atlanta airport, the Macon airport may be more convenient for some because it eliminates the need for the hour plus ride in a shuttle van. However, you may find the flight schedule to Macon requires that you plan to arrive on Thursday night before the conference instead of Friday morning of the conference.
A. General Conference Information (cont.)

Conference Hotel: The conference hotel is the Crowne Plaza hotel located at 108 First Street, Macon, Georgia, 31201, in downtown Macon. We will have a shuttle service from this downtown location to the law school, which is approximately ½ mile away from the hotel. The times for shuttle transportation are included in the conference program (included in these materials). The hotel conference rate is $79 per night plus 12% sales tax (tax rate is subject change). **The cut off date for this rate was March 7, 2006, at 3:00 pm.** Please call the hotel directly at (478) 746-1461 and mention “Critical Tax Theory Conference at Mercer University School of Law.” Check-in time is 4:00 pm and check-out time is 11:00 am. Per its cancellation policy, the hotel will charge for reservations not cancelled by 6:00 pm on day of arrival.

All hotel reimbursement requests for up to two nights in the conference hotel should have already been provided to Professor David A. Brennen via e-mail at (brennen_da@mercer.edu). If you think you have made an e-mail request for hotel reimbursement but did not receive an e-mail message in March confirming your request, please let Professor Brennen know immediately.

Additional Hotel: An additional hotel location for those who enjoy historic inns is the 1842 Inn. The 1842 Inn is a four diamond historic bed and breakfast inn located at 353 College Street, Macon, GA 31201, approximately one block from the law school. The telephone number for the 1842 Inn is (877) 452-6599. The rate is $189 plus 12% tax plus 10% environmental charge (tax rate and environmental charge are subject to change). In order to reserve a room at the 1842 Inn, you must pay for one night in advance. **The law school will not reimburse any portion of the room charge for this additional hotel.**

Meals During the Conference: Mercer University School of Law will provide several complimentary meals for conference participants. On Friday afternoon we will have a welcome reception at the law school, complete with snacks and soft drinks. On Friday evening we will have a cocktail reception and dinner at a local restaurant called the Tic Toc Room, which is famous for its “blues singers” history. On Saturday morning we will provide a continental breakfast and snacks and drinks throughout the day. On Saturday afternoon we will have a buffet lunch at the law school. On Saturday evening we will have dinner at the Back Burner Grill.

Brief History of the Critical Tax Theory Conference (1995-2006): The Critical Tax Theory series of conferences began in 1995. The principle organizer of this first conference was Professor Nancy Staudt, then on the faculty at State University of New York at Buffalo. One of Professor Staudt's Buffalo colleagues coined the term "critical tax theory" as a lark one day and the name stuck. The name was apt because it was so well-known as part of the original group of schools (along with Wisconsin) to push along critical legal studies. Plus the name seemed to fit the papers since they were mostly about tax law as it relates gender, race, and sexual orientation. Professor Staudt organized another tax conference in 1997 and invited many of the same people and called
A. General Conference Information (cont.)

it "Democracy and Taxation." This seemed to fit the papers and the focus was more on political science and tax.

In 2000, Professor Beverly Moran approached Professor Staudt about her willingness to work with Professor Moran to organize a third conference and make the conference an annual event. They named this annual event “Critical Tax Theory Conference” as a throw-back to the original conference in 1995. That conference in 2000 was held at University of Wisconsin, which is where Professor Moran was teaching at the time.

Although Professors Staudt and Moran discussed the name of the conference and what it meant, they decided to keep the name for what it stood for. They did not intend the name to in any way restrict the content of the conference. In fact, Professor Staudt intended just the opposite: she hoped the conference would develop beyond any narrow focus. Over the years, the conference has developed in this vein, with the proviso that the host of the conference has complete authority to organize the conference in whatever way she/he chooses.

The conference conveners have never defined what it means to be a critical tax theorist. Professor Beverly Moran, one of the key original organizers of the conference, persuaded the group to keep the name "critical tax theory" because it sounded cool but not necessarily because attendees only did critical theory analysis. In fact, many non-crits have attended and presented great papers from the very start in 1995.

Since 2000, the conference has been held at Washington University in St. Louis (2001); Tulane University (2002); University of Michigan (2003); Rutgers University at Newark (2004); and Seattle University (2005). Thus, this year’s conference – held at Mercer University in 2006 - marks the 9th in the series of Critical Tax Theory Conferences.
B. Conference Program

FRIDAY, APRIL 7, 2006
Mercer University School of Law, Moot Court Room

Shuttle at 12:30 at Conference Hotel (Crowne Plaza)

1:00 – 1:15 Introductory Remarks

Professor David A. Brennen, Mercer University School of Law
Dean Daisy H. Floyd, Mercer University School of Law

1:15 – 2:45 PANEL #1: Tax Administration and Enforcement

Moderator:
Leandra Lederman, Indiana University

Panelists:
Sagit Leviner, University of Michigan
A New Era of Tax Enforcement Policy - From 'Big Stick' to Responsive Regulation
Charlotte Crane, Northwestern University
The Curious Case of Judicial Restraint in the Review of Tax Systems: Some Sources in History
Lily Kahng, Seattle University
Taxing Disaster: The Administrative Domain of the IRS

2:45 – 2:50 BREAK

2:50 – 3:50 PANEL #2: Tax Accounting and Taxing Debt

Moderator:
Monica Armstrong, Mercer University

Panelists:
Neil Buchanan, Rutgers University - Newark
What is Fiscal Responsibility? Long-term Deficits, Generational Accounting, and Capital Budgeting
Joseph Dodge, Florida State University
Exploring the Treatment of Borrowing and Accruals in a Realization Income Tax

4:00 – 4:45 Keynote Address

Nina Olson, National Taxpayer Advocate
B. Conference Program (cont.)

Shuttle at 5:00 at law school
Shuttle at 6:45 at Conference Hotel

6:45 Reception (Tic Toc Room)

7:45 – 9:30 Dinner (Tic Toc Room)

Shuttle at 9:30 at Tic Toc Room to Conference Hotel
B. Conference Program (cont.)

SATURDAY, APRIL 8, 2006

Mercer University School of Law, Moot Court Room

*Shuttle at 7:45 am at Conference Hotel to Law School by 8:00 am*

8:00 – 8:30 Breakfast

8:30- 9:30 PANEL #3: Tax Fairness

Moderator:
Patricia Dilley, University of Florida

Panelists:
Anthony Infanti, University of Pittsburgh
*Homo Sacer, Homosexual: Some Thoughts on Waging Tax Guerrilla Warfare*
Karen Brown, George Washington University
*Are Blacks Too Heavily Taxed?*

9:30-9:40 Break

9:40 –11:10 PANEL #4: The Price of Conflict: War Taxation in American History

Moderator:
Ajay K. Mehrotra, Indiana University (Bloomington)

Panelists:
Steven Bank, UCLA
*The Price of Conflict: War Taxation In American History - The Civil War and Reconstruction*
Kirk Stark, UCLA
*Taxation & War Finance During The Korean And Vietnam Conflicts*
Joseph Thorndike, University of Virginia
*“To Take the Profit Out of War”: World War II and the Creation of the Modern American Tax Regime*

11:10 – 11:20 Break

11:20 – 12:10 Incubator Session (4-5 projects)

Conference participants discuss research projects that are in their early stages.
B. Conference Program (cont.)

12:15 – 1:15 pm  Lunch and Incubator Session (5-6 projects) (Woodruff House)

1:20 – 1:50  Incubator Session (2-3 projects)

1:50 – 2:00  Break

2:00 – 3:30 PANEL #5: Nonprofits and Philanthropy

Moderator:
Sarah Waldeck, Seton Hall University

Panelists:
Darryl Jones, University of Pittsburgh
The Problem Of Charitable Insider Trading
Wendy Gerzog, University of Baltimore
Alice Thomas, University of District of Columbia
Elective Share
Untapped Wealth Maximization: The Under Utilization of the
501(c)(3) Charitable Entity to Level the Playing Field in Urban
Communities of Color

3:30 – 3:40  Break

3:40-4:40 PANEL #6: International Tax

Moderator:
Tracy Kaye, Seton Hall University

Panelists:
Diane Ring, Boston College
International Tax Relations: Theory and Implications
Kim Brooks, University of British Columbia
Taxing Royalty Payments with a Developing Country Source: A
Comparison of Canada and Australia’s Tax Treaties

4:40 – 5:00  Planning and Concluding Remarks

Shuttle at 5:15 at law school
Shuttle at 7:00 at Conference Hotel

7:15 – 9:30 Dinner  (Backburner Grill)
Shuttle at 9:30 at Backburner Grill to Conference Hotel
C. List of Conference Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Address/Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monica Armstrong</td>
<td>Mercer University School of Law</td>
<td><a href="mailto:Monica.D.Armstrong@irscounsel.treas.gov">Monica.D.Armstrong@irscounsel.treas.gov</a></td>
</tr>
<tr>
<td>Steven Bank</td>
<td>UCLA School of Law</td>
<td><a href="mailto:bank@law.ucla.edu">bank@law.ucla.edu</a></td>
</tr>
<tr>
<td>David A. Brennen</td>
<td>Mercer University School of Law</td>
<td><a href="mailto:brennen_da@mercer.edu">brennen_da@mercer.edu</a></td>
</tr>
<tr>
<td>Kim Brooks</td>
<td>University of British Columbia</td>
<td><a href="mailto:brooks@law.ubc.ca">brooks@law.ubc.ca</a></td>
</tr>
<tr>
<td>Karen Brown</td>
<td>The George Washington University Law School</td>
<td><a href="mailto:karenbrown@law.gwu.edu">karenbrown@law.gwu.edu</a></td>
</tr>
<tr>
<td>Neil Buchanan</td>
<td>Rutgers School of Law – Newark Center for Law and Justice</td>
<td><a href="mailto:nbuchanan@kinoy.rutgers.edu">nbuchanan@kinoy.rutgers.edu</a></td>
</tr>
<tr>
<td>Charlotte Crane</td>
<td>Northwestern University School of Law</td>
<td><a href="mailto:ccrane@law.northwestern.edu">ccrane@law.northwestern.edu</a></td>
</tr>
<tr>
<td>Patricia Dilley</td>
<td>University of Florida College of Law</td>
<td><a href="mailto:dilley@law.ufl.edu">dilley@law.ufl.edu</a></td>
</tr>
<tr>
<td>Joseph Dodge</td>
<td>Florida State University College of Law</td>
<td><a href="mailto:jdodge@law.fsu.edu">jdodge@law.fsu.edu</a></td>
</tr>
<tr>
<td>Wendy Gerzog</td>
<td>University of Baltimore School of Law</td>
<td><a href="mailto:wgerzog@ubalt.edu">wgerzog@ubalt.edu</a></td>
</tr>
<tr>
<td>Anthony Infanti</td>
<td>University of Pittsburgh School of Law</td>
<td><a href="mailto:infanti@law.pitt.edu">infanti@law.pitt.edu</a></td>
</tr>
<tr>
<td>Darryll Jones</td>
<td>University of Pittsburgh School of Law</td>
<td><a href="mailto:jones@law.pitt.edu">jones@law.pitt.edu</a></td>
</tr>
<tr>
<td>Lily Kahng</td>
<td>Seattle University School of Law</td>
<td><a href="mailto:kahngl@seattleu.edu">kahngl@seattleu.edu</a></td>
</tr>
<tr>
<td>Tracy Kaye</td>
<td>Seton Hall University School of Law</td>
<td><a href="mailto:kayetrac@shu.edu">kayetrac@shu.edu</a></td>
</tr>
</tbody>
</table>

8
### C. List of Conference Participants (cont.)

<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
<th>Address 1</th>
<th>Address 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leandra Lederman</td>
<td>Indiana University School of Law</td>
<td>211 South Indiana Ave</td>
<td>Bloomington, IN 47405</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:llederma@indiana.edu">llederma@indiana.edu</a></td>
<td></td>
</tr>
<tr>
<td>Sagit Leviner</td>
<td>University of Michigan School of</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Law</td>
<td><a href="mailto:sleviner@umich.edu">sleviner@umich.edu</a></td>
<td></td>
</tr>
<tr>
<td>Stephen Mazza</td>
<td>University of Kansas School of Law</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:smazza@ku.edu">smazza@ku.edu</a></td>
<td></td>
</tr>
<tr>
<td>Michael McIntyre</td>
<td>Wayne State University Law School</td>
<td>471 W. Palmer St.</td>
<td>Detroit, MI 48202</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:aa0884@wayne.edu">aa0884@wayne.edu</a></td>
<td></td>
</tr>
<tr>
<td>Ajay Mehrotra</td>
<td>Indiana University School of Law</td>
<td>211 South Indiana Ave</td>
<td>Bloomington, IN 47405</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:amehrotr@indiana.edu">amehrotr@indiana.edu</a></td>
<td></td>
</tr>
<tr>
<td>Nina Olson</td>
<td>Internal Revenue Service</td>
<td>1111 Constitution Ave, NW, Room 3031</td>
<td>Washington, DC 20224</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:Nina.E.Olson@irs.gov">Nina.E.Olson@irs.gov</a></td>
<td></td>
</tr>
<tr>
<td>Diane Ring</td>
<td>Boston College of Law</td>
<td>885 Centre Street</td>
<td>Newton, MA 2459</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:diane.ring.1@bc.edu">diane.ring.1@bc.edu</a></td>
<td></td>
</tr>
<tr>
<td>Kirk Stark</td>
<td>UCLA School of Law</td>
<td>Box 951476</td>
<td>Los Angeles, CA 90095-1476</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:stark@law.ucla.edu">stark@law.ucla.edu</a></td>
<td></td>
</tr>
<tr>
<td>Alice Thomas</td>
<td>University of the District of</td>
<td>4200 Connecticut Avenue, NW</td>
<td>Washington, DC 20008</td>
</tr>
<tr>
<td></td>
<td>Columbia School of Law</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:AMTDCLaw@aol.com">AMTDCLaw@aol.com</a></td>
<td></td>
</tr>
<tr>
<td>Ann Thomas</td>
<td>New York Law School</td>
<td>New York, NY</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:athomas@nyls.edu">athomas@nyls.edu</a></td>
<td></td>
</tr>
<tr>
<td>Joe Thorndike</td>
<td>University of Virginia</td>
<td>2667 English Oaks Circle</td>
<td>Charlottesville, VA 22911</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:jthorndike@virginia.edu">jthorndike@virginia.edu</a></td>
<td></td>
</tr>
<tr>
<td>Sarah Waldeck</td>
<td>Seton Hall University School of</td>
<td>One Newark Center</td>
<td>Newark, NJ 07102</td>
</tr>
<tr>
<td></td>
<td>Law</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:waldecsa@shu.edu">waldecsa@shu.edu</a></td>
<td></td>
</tr>
<tr>
<td>Virginia Williams</td>
<td>Mercer University School of Law</td>
<td>Macon, GA 31207</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:williams_vf@mercer.edu">williams_vf@mercer.edu</a></td>
<td></td>
</tr>
</tbody>
</table>
D. Biography of Keynote Speaker

Nina E. Olson is the National Taxpayer Advocate and serves as an advocate for taxpayers to the IRS and Congress. She leads the Taxpayer Advocate Service, a nationwide organization of more than 2,000 taxpayer advocates who help U.S. taxpayers resolve problems and work with the IRS to correct systemic and procedural problems. The National Taxpayer Advocate is required by statute to submit two annual reports to Congress. The first report identifies the objectives of the Office of the Taxpayer Advocate. The second report includes a discussion of the top 20 problems that taxpayers face in their dealings with the IRS and makes administrative and legislative recommendations to protect taxpayer rights and mitigate taxpayer problems. Nina’s work as the taxpayers’ “voice at the IRS” has won national acclaim: In 2004, Accounting Today magazine named her one of its Top 100 Most Influential People on the accounting profession. In January 2005, Money magazine named her one of 12 “Class Acts of 2004.”

Nina is an attorney licensed in Virginia and North Carolina. She was the founder and Executive Director of The Community Tax Law Project, the first independent 501(c)(3) low income taxpayer clinic in the United States. The Community Tax Law Project provides free legal services to Virginia low income taxpayers in federal, state, and local tax disputes.

Prior to her appointment as the National Taxpayer Advocate in January 2001, Nina maintained a private law practice, concentrating in tax controversy representation. From 1975 until 1991, she owned and operated Accounting, Tax & Information Services, a tax planning and preparation firm in Chapel Hill, North Carolina. Nina served as the chair of the American Bar Association (ABA) Section of Taxation’s Low Income Taxpayers Committee as well as the Pro Se/Pro Bono Task Force of the ABA Tax Section's Court Procedure Committee. Nina is the 1999 recipient of both the Virginia Bar Association's Pro Bono Publico Award and the City of Richmond Bar Association's Pro Bono Award. She also received the Lewis F. Powell, Jr. Award in 2001 from the Virginia State Bar, a special recognition for her pro bono service. In addition, Nina has served as chair of the Virginia State Bar's Special Committee on Access to Legal Services.

A.B., cum laude, Bryn Mawr College; J.D., cum laude, North Carolina Central University School of Law; LL.M. (Taxation), with distinction, Georgetown University Law Center. Adjunct Professor, Georgetown University Law Center.
E. Biographies of Conference Participants


BROOKS, KIM, Assistant Professor. University of British Columbia Faculty of Law. B.A., University of Toronto; Bachelor of Law, University of British Columbia; LL.M. (Taxation), York University, Osgoode Hall Law School. Previous Teaching Appointments: Queen's University, Faculty of Law, Kingston, Ontario. Subjects: Tax, Corporate Tax, International Tax, Torts and Law & Economics. Primary Research Interests: Corporate tax, International tax and Tax policy. Private Practice: Stikeman Elliott in their Toronto and London (UK) offices, practice focused on corporate and international tax, including tax aspects of cross-border investments and transactions, financings, reorganizations and restructurings. Member: Canadian Tax Foundation, International Fiscal Association, National Tax Association, and National Legal Committee of the Women's Legal Education and Action Fund. Leadership Positions include: managing editor of the Canadian Journal of Women and the Law and Acting Director of the Centre for Feminist Legal Studies at UBC.
E. Biographies of Conference Participants (cont.)


E. Biographies of Conference Participants (cont.)


E. Biographies of Conference Participants (cont.)


E. Biographies of Conference Participants (cont.)


E. Biographies of Conference Participants (cont.)


THORNDIKE, JOSEPH J., is director of the Tax History Project at Tax Analysts and author of the "Politics of Federal Taxation" column for Tax Notes magazine. He is also a Scholar-in-Residence at the University of Virginia. Thorndike is the author of numerous publications on tax policy, both past and present, including articles for the New York Times, the Washington Times, the American University Administrative Law Review, L'Economie Politique (France), and various scholarly collections. He is the editor, with Dennis J. Ventry, Jr., of Tax Justice: The Ongoing Debate (Urban Institute Press, 2002). Thorndike received a B.A. from Williams College and an M.A. and Ph.D. from the University Of Virginia. He is currently writing a history of tax fairness and social justice during the Great Depression and World War II.


F. Conference Papers
Introduction
The operation of the federal government is heavily dependent on income taxes: In 2005, about 43 percent of federal tax revenue in the United States came from individual income taxes and another 13 percent from corporate income taxes.\(^1\) This amounts to $927 billion and $278 billion respectively,\(^2\) an increase of 14.6 percent in individual income taxes and 46.9 percent in corporate income taxes compared with fiscal year 2004.\(^3\) However, every year, the government collects billions of dollars less in tax money than it believes is owed.\(^4\) For tax year 1981, the Internal Revenue Service (IRS) estimated that noncompliance with the tax code created lost revenue of $81.5 billion, enough to pay the national deficit for that year.\(^5\) In 1986 that number increased to $95 billion – a figure representing nearly one-half of that year’s federal budget deficit.\(^6\) This number dropped again in 1987 to $84.9 billion, but it was still significant, amounting to more than 20 percent of federal tax liability for that year.\(^7\)

This difference between taxes owed and taxes collected – otherwise known as the “tax gap”\(^8\) – is not only substantial, but it has nearly tripled over the past two decades, and it still continues to grow.\(^9\) Estimates released in February 2006 indicate that the U.S. tax gap for the 2001 tax year stands at approximately $345 billion dollars,\(^10\) leading to noncompliance rate of 16.3 percent.\(^11\) Both of these numbers fall at the high end of the range of estimates provided by the IRS in the

---

* Doctor of the Science of Law (S.J.D) Candidate - University of Michigan Law School (2002-present); Master of Laws (LL.M) - University of Michigan Law School (2002); Bachelor of Laws (LL.B) - University of Haifa Law Faculty (2000).

** I am grateful for the generous financial support provided by the Center for International and Comparative Law of the University of Michigan Law School and the Ann Arbor John M. Olin Center for Law and Economics.

1 These numbers are part of the “Combined Statement of Receipts, Outlays, and Balances,” provided by the US Treasury Department in December 2005. See: http://www.fms.treas.gov/annualreport/cs2005/receipt.pdf.

2 Id.

3 Id. However, compare with 53% of individual income tax and 11% of corporate income tax supporting the federal government in 1985 (taxpayer co.228, citing IRS, 1985:21).


5 (IRS, 1983).

6 *** 

7 ***

8 According to the IRS (February, 2006) the tax gap is “the difference between what taxpayers timely and voluntarily should have paid and what they did pay.” (IR-2006-28, Feb. 14, 2006). Also see IRS document FS-2005-14 (“The tax gap measures the extent to which taxpayers do not file their tax returns and pay the correct tax on time”). However, others suggest that the tax gap is not equal to the amount of additional revenue that would be collected by stricter enforcement. According to this argument, estimations of the tax gap merely portray the wedge between economic reality and a purely legal construct called “statutory taxes.” Changing one or the other would not necessarily lead to the desired results. For example, perfect enforcement could significantly affect the economic scenario (some firms would go bankrupt, taxpayers would modify their labor decisions, prices would change and so on), so the tax base may be altered. As a result, at least in theory, net revenue might shrink. Thus, standard measures of tax gaps must be interpreted cautiously. see Luigi Alberto Franzoni, Tax Evasion and Tax Compliance, in THE ENCYCLOPEDIA OF LAW AND ECONOMICS (B. Bouckaert and G. de Geest, eds, 1998) at 3* (hereafter: Franzoni).

9 The tax gap has been growing at about 15% annual rate in the decade preceding 1981. See Jonathan Skinner and Joel Slemrod, An Economic Perspective on Tax Evasion, 38 NAT’L TAX J. 345, 345 (1985). Note, however, that estimations of the tax gap are not without controversies, and the growth rate changes (i.e., decrease), for example, when adjusted for inflation. For more see Michael J. Graets and Louis L. Wilde, The Economic of Tax Compliance: Fact and Fantasy, 38 NAT’L L. J. 355, 355 (1985).


11 Written Testimony of Commissioner of Internal Revenue, Mark Everson, before the Senate Committee on the Budget on ‘The Tax Gap and How to Solve it’ (February 15, 2006). p.3.
spring of 2005. Through enforcement activities and collection of other late payments, the IRS’s intent is to eventually close some of this gap, resulting in a still enormous net deficit of approximately $290 billion for the 2001 tax year.

Noncompliance with the tax law may occur in various ways, including taxpayers’ failure to do any of the following: (1) accurately report the tax base, (2) correctly assess tax liability, (3) timely file tax returns, or (4) promptly pay taxes due. More than 80 percent of the tax gap comes from underreporting of taxes--mostly of income tax--meaning that the taxpayer either provides a partial report of his tax base or completely fails to acknowledge an existing liability. Due to the complex nature of the tax code and regulations, noncompliance may not be exclusively intentional and can stem from a wide range of causes, including “lack of knowledge, confusion, poor record keeping, [or] differing legal interpretations....” These problems may occur because “the taxpayer is ignorant, lazy, careless…following common practices in occupational groups or workplaces, heeding incorrect advice from the IRS … taking the advice of a tax professional who recommended strategies that shade into illegality, or many other reasons.” Unfortunately, a significant amount of noncompliance is willful. According to many estimates, the largest portion of the tax gap is due to underreporting of income by small business and self-employed individuals, the fastest growing segment of taxpayers. Most of this noncompliant behavior falls under the rubric of intentional (and illegal) tax evasion.

Actions that intentionally result in tax noncompliance can be categorized into three broad groups. On the two ends are tax evasion and tax avoidance while a third group, aggressive tax planning, is somewhere in between, though still a form of tax avoidance. Despite the fact that evasion and avoidance have much in common in the economic realm, from a legal standpoint, evasion differs significantly from avoidance in that evasion is unlawful, and hence punishable. Indeed, tax

---

13 Id. Notice, however, the statement of IRS Commissioner, Mark W. Everson: “The vast majority of Americans pay their taxes accurately and are shortchanged by those who don’t pay their fair share.” Id.
14 For example, a taxpayer could fail to file a tax return, underreport legal or illegal income, overstate deductions, report the wrong type of deduction, or refuse to pay the amount of tax owed. See John S. Carroll, How Taxpayers Think about Their Taxes: Frames and Values, in WHY PEOPLE PAY TAXES (Joel Slemrod, Ed. 1992) at 43 (hereafter: Carroll, 1992); John S. Carroll, A Cognitive-Process Analysis of Taxpayer Compliance, in TAXPAYER COMPLIANCE (Vol.2: Social Science Perspective, Jeffrey A. Roth and John T. Scholz, eds., 1989) at 228 (hereafter: Carroll, 1989); Franzoni at 5*.
16 Franzoni at 5*.
19 See, for example, IRS Stratigic Plan 2005-2009, Department of Treasury, Internal Revenue Service, publication 3744 (Rev. 6-2004) catalog number 31685B p. 9 at http://www.irs.gov (“The IRS must meet the challenges posted by small business and self-employed customers. This group of taxpayers is the fastest growing segment of taxpayers and is estimated to be the largest single contributor to the federal tax gap”), and there at p.18. Specifically, current estimates of the tax gap indicate that 32% of the gap is due to underreporting of income by high income individuals, 8% by large corporations, 3% by small corporations, 23% by employment tax, 3% by estate and gift and excise tax, and 31% by other individuals. See there at p.18.
20 Id. Also see Alex Raskolnikov, Crime and Punishment in Taxation: Deceit, Deterrence, and the Self-Adjustment Penalty, (forthcoming) at p. 6. (hereafter: Raskolnikov, 2006).
21 Murphy, 2005, at **
23 Franzoni at 4*.
evasion is commonly defined as “a deliberate failure to comply with one’s tax obligation in a manner which clearly violates the law.” This could include, for example, a failure to submit tax returns or income that is received in cash and is not reported to the tax authority. In comparison, tax avoidance (also known as “tax reduction”) occurs when the taxpayer intentionally reduces his tax liability in ways that may be unintended by the legislator but is permissible by law. Avoidance is typically accomplished by constructing business transactions in ways that minimize tax liability often through exhausting favorable tax treatments such as any of the deductions, credits, and losses available in the code. This type of noncompliance gives taxpayers more freedom to vary their compliance strategy than when evasion is involved. The third group of tax noncompliance is a specific—more extreme—type of tax avoidance commonly referred to as “aggressive tax planning” (also known as “abusive tax shelters”). Taxpayers in this group seek to exploit deficiencies or uncertainty in the law so that they comply (only) with the letter of the law while ultimately undermining the policy intent or rational behind it. An example of aggressive tax planning is ***

Tax noncompliance is a pervasive problem in all societies. Whatever shape it takes, this type of disobedience creates financial costs and furthers a climate of disrespect, antagonism, and selfishness in the relationship among citizens and between them and the government. When noncompliance with the tax law is prevalent, the tax system becomes a deficient means for pursuing and implementing government goals which rely on funding by tax revenue. With tax noncompliance, effective taxes do not reflect the statutory intention, creating disturbing results such as upsetting the distribution of tax burdens and, more generally, wealth in society. For instance, when wealthy citizens have better opportunities and means to reduce their tax liability compared with other less well-off citizens, the effective taxes are likely to result in a more regressive and less equitable system than the statutory one. Abusive tax practices also jeopardize horizontal equity when there exists an unequal distribution of opportunity to reduce or eliminate tax liability and the willingness to seize it. Furthermore, in a country with fixed revenue requirements, reducing the tax liability of any given sector of taxpayers, in effect, means that

---

24 Franzoni at 1* (italic added).
25 Franzoni at 3*. ADD Slemrod.
26 Franzoni at 3-4*.
27 Raskolnikov, 2006, at p. 17.
31 Carroll, 1989, at 229, and *** (for – among citizens**)
32 Franzoni at 1*, and there at 4*.
33 Slemrod, 1992, at 1 (“it is impossible to understand the true impact of a county’s tax system by looking only at the tax base and the tax rates applied to that base. A critical intermediating factor is how the tax law is administrated and enforced. What is apparently a highly progressive tax rate structure may, in fact, be proportional, or even regressive, if taxes levied on the wealthy are not collected. What is apparently a tax base finely tuned to reflect individual differences in ability to pay may, in fact, produce a capricious distribution of tax liabilities of the tax law is selectively enforced.”).
34 (Louis Kaplow - alex.n.79,76). Also see Alex Raskolnikov, 2006, at p.14 (Claiming that the effect of collecting (or failing to collect) revenue is, ultimately, redistributive. If evasion succeeds, the evader benefits from extra wealth while if evasion is detected and the evader is forced to pay the tax, other people benefit from the funds.)
36 Franzoni at 4*. 
higher and more distortionary taxes are levied on others. All of this, in turn, produces inefficiencies as competition is impacted by the unequal distribution of the tax burdens and by economic practices motivated by tax abuse. In addition, efforts to avoid or evade taxes are, on their face, a deadweight loss to society, as are the costs of taxpayer compliance which can be as high as 10-13 percent of total tax liability for income tax and VAT in industrialized countries. These costs are not only monetary losses but may also generate resentment and weaken taxpayers’ moral conscience, possibly causing a frustration that escalates to tax disobedience as a form of protest against the government, the tax administration, or as a means of equalizing the tax burden. Moreover, existing noncompliance, especially when widespread, undermines citizens’ sense of trust and social solidarity and their inclination to pay their fair share to society, which, ultimately, furthers the problem of tax noncompliance.

No tax system can achieve perfect compliance. Moreover, reducing the tax gap is a challenging task given the magnitude and persistent levels of noncompliance. Still, due to its size as well as the pervasive consequences of noncompliance, even a small or moderate reduction in the net tax gap could yield substantial returns, including improving the government’s ability to pursue its goals and restoring some level of socio-economic and political harmony. Based on the IRS’s recent estimates, each 1 percent reduction in the net tax gap in the United States would likely yield more than $2.5 billion dollars annually. Thus, a 10 to 20 percent reduction of the net tax gap would translate into $25 billion to $50 billion or more in additional revenue annually. As a result of the growing tax gap, in recent years the IRS has taken a number of steps to bolster enforcement. The IRS budget request for fiscal year 2005 was $10.674 billion, $490 million more than the amount proposed in fiscal year 2004. Three-hundred million dollars of this increase was allocated for enforcement. This budget increase boosted the audits of high-income taxpayers who are earning $100,000 or more to 221,000 reviews in fiscal year 2005, the highest number in

37 Andreoni, Erard, and Feinstein, 1998, at 818. For example, those whose tax burdens are easier to evade, such as self-employed individuals, may end up paying a smaller share of taxes than the rest of a given population who will need to “fill up” the gap. For that reason, tax compliance can also be linked to labor market behavior (e.g., occupational choice, human capital investment, and labor supply) as the different types of occupation and the degree to which they are pursued could affect the level of potential evasion (see Andreoni, Erard, and Feinstein, 1998, at 819).
38 Franzoni at 4*.
40 Andreoni, Erard, and Feinstein, 1998, at 833; Franzoni at 16* (arguing that legislatures should thus avoid as much as possible increasing the complexity of the tax code and regulations which raise compliance costs and consequently foster evasion. Furthermore, when the tax code is very complex, taxpayers (who can afford this) may turn to tax experts who, given their expertise, have power to influence their clients’ attitudes towards tax avoidance and evasion as well as their behavior on those issues. For that matter, evidence indicates that tax preparers generally encourage compliance with regards to unequivocal items, and discourage it with regards to ambiguous ones. Klepper and Nagin (1989b), Scotchmer (1989b), Reinganum and Wilde (1991), Erard (1993) and Franzoni (1998)).
41 ***
43 ***
44 ***
45 ***
48 The remainder of the increase, $190 million, was intended to be used for reinvestments in consumer service, maintenance of existing levels of performances, and physical infrastructure consolidation. See Id.
the past ten years.\textsuperscript{49} The number of audits of all taxpayers increased to 1.2 million in 2004, a 20 percent increase from the year before.\textsuperscript{50} As a result of these steps and others, the IRS increased its enforcement revenue by nearly 40 percent from a total of $33.8 billion in 2001 to $47.3 billion in 2005.\textsuperscript{51}

The discussion above makes evident that tax enforcement is not only a procedural matter but also raises fundamental issues that run to the core of human socialization, especially matters of economic efficiency, morality, equity, sociology and psychology of human and social behavior. Accordingly, addressing the problem of tax noncompliance, though a challenging task, may be well worth the effort in the long run in order to not only increase the revenue collected but also to create a more equitable, efficient, and just public administration and society.

During the past 25 years there has been a boom in research on tax compliance.\textsuperscript{52} This paper explores the major theoretical developments in compliance as found in the tax enforcement literature. It investigates the role of economic analysis in tax policy-making, detailing the primary approaches to enforcing compliance and highlighting their key strengths and shortcomings against available compliance improvement data. A significant portion of the paper focuses on the relevance of the latest reforms in tax administration within the United States and abroad and the way in which these reforms interface with the compliance problem. For example, this paper looks into the current emphasis on improving service for and respecting the rights of taxpayers and explores the intriguing Australian motivational postures model of compliance. The focus of the paper is on personal income tax compliance, although much of the discussion is applicable to other areas of taxation. Unfortunately, there are many important issues that fall outside the scope of this paper. Most notably, this paper does not examine the literature on the underground economy, on the difficulties with collecting taxes internationally, or on business tax evasion, nor does it consider the relative advantages and disadvantages of sales taxes or VAT compared with income taxation in terms of their enforcement.

\section{1. The Problem of Compliance}

The difficulties of tax enforcement emerge, to a great extent, since the variables that define the tax base such as income, sales, or wealth are usually not “observable.”\textsuperscript{53} That is, without detailed information about the various financial transactions taxpayers are involved in and their overall financial (and other tax-related) standing, no one but the taxpayers can know their true tax liability and, therefore, whether each taxpayer is truthful and accurate in his report to the tax authorities.\textsuperscript{54}

\textsuperscript{49} Id.
\textsuperscript{51} The increase in revenue between 2001 and 2005 due to examination measures is $9.8 million and from document matching $1.5 million, amounting to a total of $11.3 million of additional tax money collected. This is compared with an increase in revenue of $2.2 million due to other forms of collection during that time. IRS Updates Tax gap Estimates, IR-2006-28, Feb.14, 2006 at: hpp://www.irs.gov/newsroom/article/0,,id=154496,00.html.
\textsuperscript{52} Prior to 1980 the empirical work on tax compliance consisted of only a few studies, based either on surveys of taxpayers attitudes or on small, idiosyncratic datasets. Over the past decade or so, empirical research on tax (especially individual income tax) compliance has blossomed (see Andreoni, Erard, and Feinstein, 1998, at 835). ADD Slemrod & Yizhaki.
\textsuperscript{53} Franzoni at 3-5* (arguing that tax evasion and avoidance and their harmful consequences may in fact be worsened by laws, or even constitutions when they are drafted as if the tax base is observable, limiting the corrective instruments available to the government, such as setting tax rates according to their degree of enforceability).
\textsuperscript{54} Franzoni at 3*. 
To a certain extent, verifying information may be obtained by means of costly audits or third party reporting such as of banks and employers.\(^{55}\) Assuming that such information can be acquired in a timely manner and is found to be accurate and coherent,\(^ {56}\) the tax base becomes “verifiable.”\(^ {57}\) In other cases, however, as when the taxpayer is involved in transactions that are beyond the reach of the tax authority and official statistics, including when income is received by way of cash transactions, the tax base is almost impossible to verify.\(^ {58}\)

Taxpayers are able to take advantage of the unobservable nature of their tax base or, in other cases make innocent mistakes, when reporting a partial or otherwise incorrect figure on their returns, in a manner that is difficult to detect.\(^ {59}\) It can also be difficult to clearly identify which taxpayer is most likely to be non-compliant. Some of the key influences for whether a taxpayer will comply, such as the perceived probability of detection and opportunities for evasion, are rather tricky to capture and to compute.\(^ {60}\) All these shortcomings make it extremely complicated to not only detect and correct non-compliant behavior but also to study it and to better understand its possible causes and influencers.\(^ {61}\) It is in this grey area that economic analysis can intervene and offer methodological guidance.

Economic analysis plays an important role in crystallizing the issue of compliance and, specifically, pinpointing those factors involved in the compliance (or lack of it) of taxpayers with the tax law.\(^ {62}\) As compliance issues are examined, analysts must simplify the many complexities of the problems under scrutiny in order to produce a framework in which the parties and relationships involved become manageable and clear so that attention can be drawn to the essential questions that need to be resolved.\(^ {63}\) Modeling tax compliance further facilitates an important process in policy-making in which the possible consequences of establishing alternative enforcement strategies can be examined and compared.\(^ {64}\) This, in turn, allows policy-makers to deliberate on and, possibly, offer viable policy alternatives to pursue.\(^ {65}\)

\(^{55}\) ***.
\(^{56}\) ***.
\(^{57}\) Franzoni at 3*.
\(^{58}\) Franzoni at 3-4*. Attempts to detect unobservable income include inferring what would be the amount of tax dollars missing due to noncompliance in the underground economy by using indirect data such as large differences between income and expenditures at the aggregate level. Unfortunately indirect evidence are often of questionable value and suffer from sever econometric shortcomings. See Frank Cowell, *Carrots and Sticks in Enforcement*, in *THE CRISIS IN TAX ADMINISTRATION* (Henry J. Aaron and Joel Slemrod, Eds., 2004) at 237 (hereafter: Cowell, 2004).

\(^{59}\) ***.

\(^{60}\) Slemrod, 1992, at 3.
\(^{61}\) For that matter, see, for example, GAO Reports of April & October, 2004 (stressing that regularly measuring compliance is critical to the IRS’s ability to reduce the tax gap. A significant part of IRS’s tax gap estimate is based on recently collected data on individual income tax reporting compliance. However, other areas of the tax gap rely on old data and outdated methodologies. The IRS does not have approved plans, with one exception (the Earned Income Tax Credit Program), to collect more current compliance data covering various components of the tax gap. Although it can be challenging to develop, data on the reasons why taxpayers do not comply with the tax laws could help IRS more effectively tailor its efforts to reduce noncompliance. The IRS has begun to capture data on the reasons for noncompliance, but it has concerns with the data. Although the IRS is developing a system intended to better capture examination data, it does not have specific plans to develop better data more generally or to ascertain the reasons for noncompliance).


\(^{63}\) Creedy, 2000, at 135-36.

\(^{64}\) Creedy, 2000, at 135.

\(^{65}\) Cowell, 2004, at 231. Creedy, 2000, at 134-35 (indicating that policy advice based on tax models can generally take two primary forms: First, it may involve presenting the detailed implications of a specified tax scheme, so that policy-
Due to the complexity, data limitations, and difficulties in obtaining a consensus regarding the appropriate research methodology to apply, it is generally understood that no one economic model can offer a complete picture of the tax compliance problem.\textsuperscript{66} Furthermore, notwithstanding the valuable information that becomes available through application of economic models, economic analysis provides, at best, “a tentative guidance in well-defined circumstances.”\textsuperscript{67} Thinking in terms of models forces analysts to become explicit about their simplifications in terms, for instance, of what is being considered in the model, what is ignored, and the type of assumptions that are taken into account.\textsuperscript{68} Models simplify a much more complex reality making them, almost by definition, unrealistic and, therefore, subject to criticism on these very grounds.\textsuperscript{69} To a certain extent, improvements in data and in methodology help bridge the gap between reality and the more theoretically based models.\textsuperscript{70} All models, however, have their shortcomings, and these must be recognized when simulating policy scenarios and generating recommendations.\textsuperscript{71} In the end, alternative models of tax compliance may help provide a powerful insight into the different aspects of enforcement.\textsuperscript{72}

\begin{footnotesize}
\begin{itemize}
  \item \textsuperscript{66} Cowell, 2004, at 231; Creedy, 2000, at 156-57.
  \item \textsuperscript{67} Creedy, 2000, at 135-36; Cowell, 2004, at 231.
  \item \textsuperscript{68} Creedy, 2000, at 135.
  \item \textsuperscript{69} Creedy, 2000, at 136.
  \item \textsuperscript{70} Creedy, 2000, at 157. Andreoni, Erard, and Feinstein argue that empirical work on tax compliance is still young, with many of the most important behavioral hypotheses and policy questions yet to be adequately investigated.
  \item \textsuperscript{71} Creedy, 2000, at 156-57. For example, a particular difficulty with modeling tax enforcement is the specification of the variable characterizing the probability of audit. Usually some proxy for evasion opportunity (such as the presence of business income) has been used to categorize audit classes, and the probability of audit is expected to differ across these classes. However, individuals’ opportunities for participating in evasion differ greatly among occupations and social groups and may not be adequately captured in such general proxies. Similarly, specifying an empirical model can have underlying problems resulting from sample selection bias in which, for example, relevant taxpayers are entirely excluded from the model because they do not file tax forms and are therefore operating almost as ghosts in the tax system (see Cowell, 2004, at 236).
  \item \textsuperscript{72} As recently pointed out by John Creedy, 2000, at 134: “…Economic models come in all shapes and sizes, depending on the nature of the policy issues examined.” Creedy also noted: “Important lessons can be learned from the use of relatively small, although often sophisticated, tax models … However, direct policy advice often requires the use of microsimulation models that are apt to reflect the kind of population heterogeneity found in real life situations” (Creedy, 2000, at 156). Also see Creedy, 2000, at 137, and there at 156-57, and Cowell, 2004, at 231.
\end{itemize}
\end{footnotesize}
The Curious Case of Judicial Restraint in the Review of Tax Systems: 
Some Sources in History 
Charlotte Crane¹

Contemporary courts often have seemed reluctant to interfere with the tax systems of the governments they oversee. This reluctance takes many forms, all of which result in a relatively limited judicial role in tax administration. Some of this reluctance shows up when courts give relatively narrow readings to the assignments they have been given by the legislative branch.² Sometimes it shows up when courts are asked to consider aspects of tax administration that may raise constitutional concerns. Thus, on several occasions in the last few decades, the Supreme Court of the United States has determined that Congress could not have intended that state or federal courts be required to issue orders interfering with state tax systems as a result of federal statutes allowing general challenges to state violations of civil rights, so long as the states themselves provided adequate remedies for aggrieved taxpayers.³ The reluctance also shows in the way

¹ Professor of Law, Northwestern University School of Law.

² For instance, the Supreme Court has never repudiated the holding in Felt v. Tarrant, 283 U.S. 269 (1931)(dismissing suit for recovery of improperly computed excess profits tax for failure to specify the grounds for refund in the administrative refund claim) that, even when the government acknowledges that an erroneous payment has been made, relief cannot be granted without compliance with the conditions set by Congress for asserting claims against the United States. Felt is a particularly harsh precedent, given that the court acknowledged that futility related to the requirements would not justify deviation from them. It is nevertheless relied upon frequently by the lower federal courts, e.g., Bartley v. United States, 123 F.3d 466 (7th Cir. 1997) cert. denied, 522 U.S. 1062 (1998) (dismissing clearly frivolous class action challenging the “continuing over-collection of federal taxes”).

³ National Private Truck Council v. Oklahoma Tax Comm’n, 515 U.S. 582 (1995)(federal civil rights statute could not be used as basis of suit to enjoin taxes claimed to be unconstitutional when adequate state remedies available, and therefore attorneys fees available under that statute not available); Fair Assessment in Real Estate v. McNary, 454 U.S. 100 (1981)(substantially the same limit on federal courts, resting on grounds of intergovernmental comity).

Hibbs v. Winn, 542 U.S. 88 (2004) may prove to be a substantial departure from the behaviors described in the text. In Hibbs, a deeply divided court held that the Tax Injunction Act did not bar a suit challenging on
various doctrines relating to the judiciary’s own power—those relating, for instance, to standing, class actions, and the appropriateness of injunctive relief. Thus the federal courts continue to decline to allow challenges to the administration of the federal income tax to proceed except in the context of litigation involving an actual liability of a taxpayer through proceedings specifically prescribed by legislation. The lower federal courts have consistently followed the Court’s lead, and have held that a taxpayer’s remedies against the tax collection efforts of the federal government are limited to those provided by federal statute, and that any other result would be inconsistent with the limited waiver of federal sovereign immunity those statutes represent.

first amendment grounds a taxpayer challenge to a tax credit. If one reads Hibbs as removing the TIA bar from all challenges to state tax regimes that do not threaten collections, rather than only those that involve both a first amendment challenge and no threat to collections, there will be increased pressure on notions of taxpayer standing, and thus an increased likelihood that the backdrop of which NPTC and FAIRE are part will change. See, e.g., Cuno v. Daimler Chrysler, Nos. 04-1704 and 04-1724

E.g., Allen v. Wright, 468 U.S. 737 (1984)(denying standing to challenge the tax exemption of racially segregated schools); Bob Jones University v. Smith, 416 U.S. 725 (1974) (same); but see South Carolina v. Regan, 465 U.S. 367 (1983)(allowing an original suit to challenge as unconstitutional a statute conditioning the federal income tax exemption for state-issued bonds on the registration of the bonds); Commissioner v. Shapiro, 424 U.S. 614 (1976)(holding that the federal Tax Injunction Act, 26 U.S. 7421 did not bar a suit for an injunction against a jeopardy assessment levied upon the taxpayer’s extradition)

In general, the need to prove, on behalf of each taxpayer claiming a right to a refund, satisfaction of the technical requirements predicate to a waiver of sovereign immunity, have been thought to preclude class certification. Exceptions do exist. In Appoloni v. United States, 218 F.R.D. 556, modified by 219 F.R.D. 116 (W.D. Mich. 2003) (in case challenging the application of FICA taxes to lump-sum early retirement payments made in circumstances in which the plaintiffs gave up tenure rights, where plaintiff’s counsel had documentation that all class members had complied with all prerequisites); Klender v. United States, 218 F.R.D. 161, modified by 218 F.R.D. 551 (E.D. 2003) (involving essentially the same claims and the same procedural posture). Neither decision amounts to approval of the sort of open-ended class actions typical of state law mass tort actions.

The general rule, in both the district courts and in the Court of Federal Claims, remains hostile to class actions. E.g., Fisher. v. United States), (No. 04-1726T) (Jan. 6, 2006 Court of Federal Claims)(denying certification without prejudice, even after acknowledging that new court rules render inappropriate the general assertion that, given the sovereign immunity interests of the United States, class actions are disfavored in that court)

Taxpayers (and their counsel) are eagerly awaiting the fate of a recent class action filing by Baker and McKenzie on behalf of Radio Shack, seeking class certification of a suit challenging the validity of the telephone excise tax, and asserting up to $9 billion in refund claims, in the wake of successful suits brought by large individual taxpayers challenging the tax and the government’s continued persistence in collecting the tax. See Crystal Tandon, Telephone Tax Challenged in Class-Action Suit, 2006 TNT 7-3 (January 11, 2006).
These decisions reflect but one piece of the armor that appears to protect most tax systems, both federal and state, from challenge and disruption at the hands of the courts. The armor consists both of statutory prohibitions against pursuing such challenges outside of narrowly prescribed routes, and in doctrines dictating judicial restraint even where no such prohibitions might apply. It includes both doctrines that limit relief prior to making tax payments and that limit the type of challenge that can be made after payment; it also doctrines that restrain state courts with respect to state taxes, and federal courts with respect to federal and state taxes. (This approach no longer holds with respect to some state court systems, which have accepted innovations in procedures relating to taxes, including class actions.\(^6\))

The result of these doctrines is a procedural backdrop that is starkly different from that through which most other bodies of public law are developed.\(^7\) Most other bodies of public law, especially the federal administrative law with which American legal academics (and hence many practicing lawyers) are most familiar, are the product of legislative and administrative processes that are constantly subject to judicial review. This review extends to the legislation from which the administrative agency derives its power, the general administrative rules as they are developed by the agency, and to their application in individual cases. Frequently, this review is available to anyone with a plausible claim that they will be concretely affected by the emerging law, and frequently as an alternative to continued pursuit of administrative procedures. This more generally available review is part and parcel of the almost pervasive view that virtually all governmental action, to the extent that it affects the particular rights of particular individuals, should be subject to examination by the courts. Many of these procedures,

---


\(^7\) Compare, e.g., General Electric Co. v. Environmental Protection Agency (DC Cir. March 2, 2004)(allowing judicial review of the administration of CERCLA despite apparent attempt to allow agency to proceed without judicial interference).
however, are not available with respect to tax administration, particularly federal tax administration. Put another way, in many respects, no government actor is less likely to be accountable to a court for the rules and procedures he fashions than the Commissioner of the Internal Revenue Service.

As a result of this armor, the bulk of substantive tax law rules are never subject to review by courts at all.\(^8\) This phenomenon is currently starkest in the federal income tax system. In most situations, only a taxpayer who is actually liable can challenge the substantive rule on its face, and then only after he has actually been billed for a deficiency, and sometimes even paid a tax.\(^9\) And such a challenge is likely to be entertained by the courts only if it is pursued through the routes carefully prescribed by legislation, which is likely to include renewed administrative, rather than judicial, procedures. Even after payment and adherence to the prescribed route for challenge, such a taxpayer is likely to be able to obtain a judicial hearing of her challenge only if her administrative refund claim is denied. Thus, often only those tax rules that result in the continued assertion by the

---

\(^8\) This phenomenon was publicized rather broadly as a result of the litigation in Commissioner v. Warren, 282 F.3d 1119 (9th Cir. 2002) over the constitutionality of the exemption for the “parsonage allowance” available to “ministers of the gospel” under section 107 of the Internal Revenue Code. The case was treated by the court as mooted by an act of Congress, which appeared to be motivated by a desire to remove the case from the courts.

\(^9\) This ability to control which claims are in fact asserted means that those that might count on the federal tax consequences of a transaction cannot proceed, or can proceed only on relatively unfavorable terms, given their inability to remove uncertainties regarding tax consequences. In one recent case, a federal district court overseeing a bankruptcy proceeding held that it had no jurisdiction to determine the tax issues that would establish the value of the bankrupt’s estate so long as the IRS had not challenged the bankrupt’s claims for cost recovery and fuel credits, even though these claims allegedly had a significant impact on the value of the bankrupt estate and thus on the viability of the confirmed plan. The value of the tax benefits at issue depended upon the time at which assets related to various processes were in fact placed in service. The IRS had not participated in the bankruptcy proceeding, although issues related to those in question in the litigation would have affected the bankrupt estate’s tax liabilities and thus the outcome of the proceeding in bankruptcy. While the bankruptcy proceedings were pending, the IRS had issued a private letter ruling suggesting that the credit might not be available (and in the process of making this suggesting indicating that it might not consider itself bound by the factual assumptions underlying the confirmation order), thus reducing substantially the yield that the bankrupt estate was able to enjoy through its efforts to “monetize” the credit. In re Dycoal, No. 99-24594, Bankr. W.D. Pa., June 30, 2005), aff’d., W.D. Pa, No. 05-679 Feb. 15, 2006.

[Blackfeet Bank insurance cases; cf Australian practice]
enforcing agency of liability of a taxpayer are reviewed. Those rules that result in no liability, because taxpayer behavior has been altered to avoid the tax, are not subject to review. Nor are those rules that the enforcing agency simply chooses not to assert once taxpayers resist ever subject to challenge.

This hands-off approach to tax systems is manifested in many ways, all of which have a tendency to further the gap between the judicial role in tax matters and that in other areas of administrative law. When courts assume as a limited jurisdiction over a subject matter as they frequently have in tax matters, it seems logical that they will invest less in developing a reputation in dealing with that subject matter, and thus may be less inclined to question the government’s position or to attempt to understand how the limited issue before it fits in the larger regulatory regime. The two phenomena are likely to interact in such a way that the court’s role becomes less and less useful under any view of the purpose of judicial participation. Even when review is undertaken, the reviewing court is not likely to understand the ramifications of the result reached, either in terms of the evolution of the substantive law, or in terms of the overall administrability of the system. Neither the tax administrators nor the public are likely to see the courts as of much use in promoting the overall legitimacy of the tax administration system, or even, perhaps, as useful adjudicators in individual disputes. Thus, what might at first glimpse seem to be a judicial inclination to protect tax systems, may in fact undercut them in the long run.

What is the source of this reluctance of current courts, particularly the federal courts, to take for themselves a larger role in tax administration and collection? Why not allow class actions? Why not find non-injunctive remedies inadequate more often? Why construe jurisdictional statutes (and the waiver of sovereign immunity inherent in them) so narrowly? From a contemporary perspective, some of the more important factors seem to relate to the judge’s own views of the nature of courts and judicial relief. Indeed, much of the current attitude toward challenges to taxes seems to have changed very little since the time when courts were uncomfortable with declaratory judgment actions, much less with class actions and other modern judicial devices that make challenges in judicial fora more appealing than challenges in the legislature. Limitations on the judicial power over the public fisc have always had considerable appeal to judges
when to act without such limits would be to ask them to consider the legitimacy of tax provisions. Threats to the legitimacy of enacted tax provisions tend to disrupt budgets and destabilize politics. Budgeting and resource allocation are quintessentially legislative functions, and courts frequently display a reluctance to interfere with them for that reason alone.

This sense of limit on judicial power thus may be closely linked with traditional notions both of sovereign immunity and of separation of powers. This attitude has not been greatly affected by an apparent trend away from treating sovereign immunity as an absolute bar on jurisdiction, and toward treating sovereign immunity as a part of the substantive rules that are within a court's competence to invoke in fashioning substantive rules.

On close examination, it appears that these contemporary attitudes are not simply the result of the pragmatic instincts of the modern judiciary. There is within the tradition of many jurisdictions in the United States, a surprisingly strong historical tradition in which the judiciary plays only very limited roles in tax administration. In this paper, I want to explore some of the possible historical bases for these views of the relationship between the power to tax and the judicial power. Although some of this history fits easily within traditional frameworks of sovereign immunity and separation of powers, some of it suggests that tax proceedings created special challenges to those charged with giving meaning to separation of powers in the years immediately following the adoption of the first state constitutions and the federal constitution.

[Note to the reader: what follows outlines only one strand of judicial interaction with tax procedures: the availability of the courts to assist in collection]

In the late eighteenth century, tax administration was a relatively freestanding governmental function. (Note that in this era, separation of powers was only an asserted goal, without much particular content, and just as likely to conjure up notions of incompatibility of offices held by particular individuals as notions of limitations on the autonomy of those offices.) Most tax collectors (like many other municipal officers) were relatively autonomous, but only to the extent of the specific powers which they had
been granted. Rarely were the powers given to any municipal officer anything like
general police powers. A tax collector would be issued a warrant by the selectmen or
other body charged with setting individual tax obligations. The warrant authorized and
required him to collect specific amounts under a legislative act imposing the particular
tax. Frequently he personally could be held accountable for any taxes unpaid, and
frequently his authority to collect these taxes would remain until they all had been
collected, even if other collectors were charged with collecting taxes on the same
property under subsequent legislative acts. He was free to seize property and sell it to
satisfy his obligation under his warrant, usually without the need to enlist the approval of
a court or the assistance of other officers.¹⁰

On the other hand, he often would be personally liable for any “unwarranted” actions. A
gradual evolution from collector as independent agent acting under warrant to collector as
bureaucrat appears to have occurred over the course of the nineteenth century. More
protection from individual responsibility was provided for the collector, who over the
course of the century was less and less likely to be held accountable for all taxes included
in his warrant, so long as timely return of delinquencies was made, or to be accountable if
the warrant issued to him required him to collect a tax later held to be illegal or required
him to collect taxes from a taxpayer who was in fact exempt. But with this increased
protection, he was much more likely to be required to obtain judicial approval before
acting in a summary fashion against a taxpayer and was less and less likely to be viewed
by the courts as having a vested interest in the share of taxes granted him as a
commission.¹¹

¹⁰ Indeed, this lack of judicial or other remedy outside of the tax system itself would sometimes serve as a
justification for judicial intervention in other limited circumstances. In Preston v. Boston, 29 Mass. 7, 14,
12 Met. Xx, Judge Shaw observed that “the warrant to a collector, under our statute for the assessment and
collection of taxes, is in the nature of an execution, running against the person and property of the party,
upon which he has no day in court, no opportunity to plead and offer proof, and have a judicial decision of
the question of his liability,” and therefore concluded that one who was subject to no tax at all could sue in
assumpsit when he had paid while asserting his exemption, even though one who was only overtaxed
would only have remedies against the municipal body setting the tax.

¹¹ Compare Hinchman v. Morris, 29 W. Va. 673 (1887); McIrney v. Reed, 23 Iowa 410 (18XX) (holding
that a collector had no right of action against those on whose behalf he had earlier paid taxes, since to allow
Courts—at least as we think of them—were involved in tax administration in surprisingly limited ways in late colonial period, and their involvement in the bureaucratization of tax collection was in many ways remarkably slow. Although bodies called “courts” set tax rates for counties in many colonies, these bodies were likely to be all-purpose county-wide administrative agencies. Their dockets including determining which public works would be funded (and sometimes doing so at the behest of a private party claiming that some other municipal body was neglecting its duty), what rate would be applied, whether tax exemptions should be available; more often than not, these bodies were acting in ways that we would describe as legislative or executive, rather than judicial. Their decisions would be blatantly made on the basis of pragmatic policy, with the opinions of many diverse parties taken into account, rather than on the basis of evidence and precedent invoked by individuals who would be immediately affected by any judgment rendered. Only in relatively limited circumstances would such bodies be availed of by individuals in proceedings that might begin to be recognized as judicial, for instance, when under the applicable statutory provisions they were permitted to direct challenging appraisals.

At least in some jurisdictions, courts—as resolvers of disputes on the basis of evidence presented to them—were simply not expected to be directly involved in the messy business of actually collecting taxes. Indeed, there seems to have been little need for them to be involved in this process. Under the typical approach to property taxes, once the tax list was drawn up, and the rates as duly determined by the town meeting, county court and colonial/state assembly applied, a warrant would issue authorizing the collection of the tax.\(^\text{12}\) This warrant brought with it adequate powers of distraint—the such a suit would be tantamount to allowing tax farming: “It would not do to so hold, that a city could delegate or farm out either its taxing power or its power to enforce the collection of taxes. It would be a startling proposition to affirm, that a city could, for example, sell and assign its tax-list to any person and authorize him to exercise the high and delicate powers conferred upon the corporation. Why not? The legal answer is, that these powers are conferred upon the municipality to be exercised by it, not to be delegated to others.” [citing Thompson v. Schermerhorn, 2 Seld. 92] with Packard v. Tisdale, 50 Me. 376 (1862)(same result, but with an overtone that the collector had failed to do his duty in not proceeding to sell the taxpayer’s property for non-payment as he had initially taken steps to do.)

\(^\text{12}\) Under the existing procedures for most general property taxes, the assessor’s list was the basis on which a warrant or bill for taxes would be rendered. This warrant was the basis for the actions of the tax collecting
power to seize and sell the defaulting taxpayer’s property—similar to (but not necessarily the same as) the power that would be enjoyed by a county sheriff if given a warrant to execute a judgment on a debt. Ordinarily, the collector’s ability to distraint involved the right to sell only the taxpayer’s personal property, not his real property, but may also have included the right to arrest and imprison the offending taxpayer. In the most common circumstances, there would be no reason for the government to invoke the powers of a court to assist in tax collection, since ultimately the collection techniques available would be similar. (Indeed, invoking the courts would be unattractive to the bureaucracy, and had a force equivalent to a legal judgment, in that it ordinarily empowered the collector to take direct measures against the taxpayer without any additional official order. See, e.g., Samuel Freeman, The Town Officer, or the Power and Duty of Selectman … and other Town Officers 63-64 (Boston 1793) (Evans No. 25512).

In many jurisdictions, the sessions or county court could play a role if a taxpayer failed to file a list, see e.g., Baldwin v. Hewitt, 88 Ky 673 (1889)(clearly distinguishing this body in this role from judicial participation in tax collection).

13 This limitation remained in many instances until relatively late in the nineteenth century. See, e.g., Board of Education of Cabin Creek v. Old Dominion, 18 W. Va. 441 (1881)(noting lien on land for school taxes, and collector’s ability to sell taxes, provided by 1869 legislation)

14 See, e.g., Hibbard v. Clark, 56 N.H. 155 (1875)(“By statutes of this state ample and severe measures are provided for the collection of taxes. The real estate of the person assessed is holden for one year from the first day of June following his assessment; his goods and chattels, with limited exceptions, are liable to distraint; and for want thereof his body can be taken and committed to the common jail.”); Shaw v. Peckett, 26 Vt. 482 (1854)(successful suit for false imprisonment when collector refused to release after tender of full amount of tax and costs, but not interest, since taxes are not debts on which interest accrues; note that this suit likely to be arrest and suit in form deliberately taken as test suit).

15 The constitutionality of such summary proceedings against claims that due process is not provided where no officer vested with “the judicial power” has intervened was upheld in Murray’s Lessee v. Hoboken Land & Improvement Co, 59 U.S. (18 How.) 272 (1856)(upholding the title of one who took by sale after distraint, based upon a warrant issued by the Solicitor of the Treasury under the authority of the Act of May 15, 1820, 3 Stat. 592, providing summary remedies against defaulting customs collectors. The Court observed that “there has been no period, since the establishment of the English monarchy, when they has not been, by the law of the land, a summary method for the recovery of debts due to the Crown, and especially those due from receivers of the revenues. ...It is certain that this diversity in “the law of the land” between public defaulters and ordinary debtors was understood in this country, and entered into the legislation of the colonies and provinces, and more especially of the States, after the declaration of independence and before the formation of the constitution of the United States. Not only was the process of distress in nearly or quite universal use for the collection of taxes, but what was generally termed a warrant of distress, ... was issued to some public officer, to whom was committed the power to ascertain the amount of the default, and by such warrant proceed to collect it.” Justice Curtis then cited at least 5 early state provisions, as well as provisions implementing both the 1798 and the 1813/1815 federal land taxes describing equally summary proceedings..
extent that these courts and their agents might expect even other municipal actors to pay fees.)

There is evidence that in some circumstances, collectors did unsuccessfully seek the assistance of courts. In a 1763 case, Ruddock v. Gordon, the courts of common pleas in Massachusetts refused to entertain suits by collectors against individual taxpayers. At least one historian has suggested that the courts’ refusal to entertain the suit marked a clear change in the courts’ practice. But regardless of whether it represents a break from prior practice, or merely a strong restatement of it, the case raises as yet unanswered questions about, first, why the collector thought the courts could help, and second, why the courts refused to help. Further investigation may reveal that there was a technical legal reason for invoking the courts--for instance, if the taxpayer held real property in the collector’s jurisdiction but held no distrainable personal property there, the ability to sue in court may well have expanded the physical jurisdiction in which execution could occur. But there may also have been a more pragmatic reason—the collector’s summary remedies, even if otherwise parallel to the remedies available through judicial proceedings, would be undertaken by the collector himself or his close associates. By proceeding in court, the collector could not only forestall the unpleasant consequences without being accused of being lax, but also dissociate himself from the physical acts

The statute at issue in Murray’s Lessee did, as was typical, allow the collector to come into court to challenge the distraint and his underlying obligation. Commentators disagree about how important this provision was to the Court’s conclusion. Compare James Pfander, Article I Tribunals, Article III Courts, and the Judicial Power of the United States, 118 Harv. L. Rev. 643 (2004) with Caleb Nelson, Adjudication in the Political Branches, Feb. 15, 2006 draft on file, p. 36 n. 101.

[Note that the precedents invoked here related to summary process against the collector, not necessarily against the taxpayer.


17 Indeed, it is possible that the collector would have no remedy against a taxpayer’s real property unless the action for debt, with the possibility of execution against real property, was available.
involved in collection. The sheriff and the sheriff’s associates would be seen seizing the delinquent’s goods and selling them, not the collector himself.

The Massachusetts assembly reacted to the court’s decision by providing the collectors with a right to sue.\textsuperscript{18} But the provision was subject to sunset when enacted, and was allowed to expire two years later. It was not reinstated until much later in the century.\textsuperscript{19} Two interpretations of these later developments are possible: either it was so obvious that the collector could enlist the aid of the courts that further legislation was not thought necessary, or a conscious decision was made to deny the collector the use of the courts. There is considerable evidence that the latter interpretation prevailed for many purposes, even if it was erroneous when first made. Many state courts\textsuperscript{20} and treatises\textsuperscript{21} reporting

\begin{footnotesize}
\textsuperscript{18} 1763 acts ch. 18.

\textsuperscript{19} A 1785 statute authorized an action at law only in very limited circumstances; when a taxpayer died or removed from a town, or when an unmarried woman married without having paid the tax. In 1859 a more general right to sue if the tax remain unpaid for more than one year was introduced; in 1889 the time was reduced to three months. During the course of the nineteenth century, the longstanding prohibitions on the sale of real property for unpaid taxes were removed, with the general increase in the likelihood that the collector would find his summary remedies, including seizure and sale, not only adequate but more desirable. See generally, Philip Nichols, Taxation in Massachusetts, page 381-82 (3rd ed. 1938).

\textsuperscript{20} Crapo v. Stetson, 49 Mass. 393, 6 Met. 106 (1844)(“It is well settled, that the law gives no remedy for the collection of taxes other than those provided by statute; and unless the mode now sought to be enforced is given by statute, it does not exist.) In Crapo, the collector faced a peculiar catch-22: he brought an action in assumpsit under a statute that would allow suit only if the taxpayer had lived in the jurisdiction when the taxes were assessed, but removed from it before the tax could be collected, but the taxpayer claimed the tax was not owed because he was not a resident of the taxing jurisdiction. Massachusetts courts also refused to allow a tax to be claimed as a setoff against a plaintiff’s claim in contract, Peirce v. City of Boston, 44 Mass (3 Met.) 520 (1842) In Peirce, the court argued that the assessment of tax was equivalent to a judgment and should be honored by the court; the court held that the statute allowing offsets in suits for debts could not be read to include such obligations as “debts.”

Other jurisdictions seem to have held firm to the Ruddock result, without qualification, on the ground that tax collection simply was no business of the courts unless the legislature made it so. Baldwin v. Hewitt, 88 Ky. 673, 6xx (1889) (refusing to aid in enforcement of personal property tax, even when no property thought to remain in the jurisdiction: “The exercise of the power of taxation is legislative, and not judicial in character. Neither the levy or collection of taxes is an inherent power of the judiciary. It has ordinarily no more power to collect them than it has to levy them. These are not judicial acts. The power to provide for the public welfare, by means of the sovereign power of taxation, has, under the distribution of governmental powers in this country, been given to the legislature. Taxes can, therefore, be collected only through legislative authority. If it does not exist, the remedy is an appeal to the law-making power for additional legislation.”)

Kentucky appears to have remained committed to this approach even when other remedies were clearly unavailable. Louisville Water Company v. Commonwealth, 89 Ky 244 (1889)( refusing to assist in the
collection of a property tax against a municipal water company, even though the collector's traditional remedies of distraint and sale would be unavailing because of legislation forbidding the sale of the property of such entities: “The exercise of the power of taxation is legislative in character, while the collection of taxes, when once authorized by the law-making power, is ministerial. The one is legislative and the other executive. Neither is a judicial act, and one department of the government should be careful not to encroach upon the domain of another. It is true that the judiciary may be called upon by the Legislature to enforce the collection of taxes in a judicial way, but it has not done so in this state, save as to railroads…, and this exceptional case inerentially says that this remedy can not be resorted to in other cases.”) [but, allowing injunction against collection of tax, Louisville Water Company v. Hamilton, 81 Ky. 517, check local politics, check for subsequent grant ];

Hibbard v. Clark, 56 N.H. 155 (1875) (holding that taxes cannot be set off against debt owed to taxpayer by town:” Judgments are the judicial sentences of courts rendered in causes within their jurisdiction. Taxes are the proportional and reasonable assessments imposed by authority of law upon the inhabitants of the state. …The imposition and collection of them are ministerial acts, and are the proper subjects of inquiry as to the manner of their assessment and the mode of their enforcement in the judicial forum; and for their collection no right of action is given);

Board of Education of Cabin Creek v. Old Dominion I.M.& M. Co, 18 W. Va. 441 (1881)( holding board had no ability to bring suit to collect tax, even where related body clearly given authority to sue for debts, since “taxes are not debts”).

Van de Griff v. Haynie, 28 Ark. 270 (1871)(in dicta : “Unless the power is specifically delegated or expressed, no right of action exists for taxes, and they cannot be turned into judgments. Both the state and municipal corporations have a much better and more expeditious remedy. They have the right by summary process, to enforce collection, by levy and sale, and when this power exists, complete and ample as it most assuredly is, it would be monstrous, without plain and express authority to that effect, to say that they could abandon, at pleasure, the usual and simple manner of making collections, and resort to judicial power for their enforcement.”

In New Jersey, on the other hand, the court appears to have concluded that a municipal corporation could not use distraint and sale to collect its taxes unless specifically authorized to do so by its charter; such “when a tax is assessed [by such corporations, each individual’s share]… becomes a debt, and, if not paid, must be recovered by the corporation in the due course of law.” Bergen v. Clarkson, 6 N.J.L. 352, 365-66 (1796). This position seems clearly to be more about distrust of the corporate form than about the role of courts in administering public finance, for the court concluded, “a body of this kind can make no law which …tends to despoil the citizen of his birthright, unless such power is actually and expressly given them by charter.” Maryland courts seem also to have been comfortable with the idea that municipal corporations were entitled to invoke them in tax collecting. E.g., Dugan v. Mayor of Baltimore, 1 G.& J. 499 (1829); Mayor of Baltimore v. Howard, 6 H.& J. 383 (1825).

Increase in judicial involvement appears to have corresponded with increase in availability of sale of land, rather than mere sale of personal property owned by taxpayer, as a remedy for collection of taxes. See, e.g., McInery v. Reed, 23 Iowa 410 (1867)(per Dillon, holding that when city charter gave city power “to levy and collect” a special tax, which by the charter was to become “a lien upon the real estate upon which it may be assessed,” the city would have the power to ask a court of equity to enforce the lien, but under the circumstances of that case, the purchaser of the land could not, after having paid the tax, sue the former owner and original taxpayer for the taxes).

In his manuscript reports of cases presented to him while Chief Justice of the New Hampshire Supreme Court, Jeremiah Smith noted that the “law authorizes distress to be made, in the case of nonpayment, on any personal estate of the person taxes. Such a tax is a personal duty authorized by law; and an action would doubtless lie to enforce payment under certain circumstances.” The editor of the 1879 publication of
state law throughout the nineteenth century dutifully declare that collectors could not use the courts unless specifically granted the authority to do so by the legislature of the state. These collectors, presumably, continued to collect through distraint and sale of the delinquent’s property.

This does not mean that courts played no role in the administration of taxes in those states that adhered to the Ruddock rule that courts could not be used by tax collectors to assist them in tax collection. Indeed, taxpayers frequently availed themselves of the courts in order to challenge the actions of the collectors, generally by asserting that the collector’s distraint was a trespass or otherwise gave rise to an action at law. But the courts were

these notes, however, included a note that “It is now understood that a suit cannot be maintained in this State to enforce payment of a tax.” Jeremiah Smith, Decisions of the Superior and Supreme Courts of New Hampshire, page 522 n.1 (“it is now understood that a suit cannot be maintained in this State to enforce the payment of a tax.”).

In some jurisdictions, the Ruddock result was interpreted as conditioned upon the provision of other remedies by the legislature, or as establishing a requirement that the tax collector exhaust the remedies specifically provided before invoking more general ones. Thomas Cooley reports that debt actions brought by the government were not common as a means of tax collection. Treatise on the Law of Taxation (2d edition 1876).

In other treatises, even the limited proposition was considerably watered down. Dillon, at section 653, acknowledges the Ruddock position, but concludes that the provision of other remedies by the legislature should not be interpreted as preempting other remedies unless a clear intent to do so is indicated. Iowa adopted this position, City of Dubuque v. Illinois Central Railroad Co., 39 Iowa 56 (1874).

22 See, e.g., Henderson v. Brown, 1 Cai. R. 92 (1803)(action in trespass against collector who collected the 1798 federal direct tax based on an erroneous assessment of the plaintiff’s property as a dwelling); Martin v. Mansfield, 3 Mass. (Tyng) 419 (1807) (challenging in a trespass action the distraint of a chaise to satisfy taxes that would be owing only if he were an inhabitant of the town); Suydam and Wyckoff v. Keys, 13 Johns. 444 (1818)(action in trover by nonresident property owners against collectors of local school tax when authority for tax was claimed to extend only to resident property owners); Henry Thaxter v. Solomon Jones, 4 Mass. (Tyng) 570 (1808)(trespass action for recovery of cow seized to pay church tax where plaintiff successfully contended he had removed to a different parish.

Suits in trespass and similar actions challenging illegal distraints remained common even after more elaborate procedures for challenging assessments and obtaining refunds were legislatively specified. See, e.g., Erskine v. Hohnbach, 81 U.S. 613 (1871)(suit against collector in trespass could proceed even in the absence of proof that taxpayer availed himself of an administrative appeal (which would have been required were the suit simply for refund), and, although collector would not under more modern law be liable even if taxpayer were not liable for tax, if process resulting in authorizing order was regular on its face, verdict against collector affirmed where evidence regarding collector’s authority presented to jury). After Erskine, however, it became clear that such an action against the assessor could not be used to
only involved in the rare case in which the taxpayer viewed a suit against the collector as worthwhile, given the relatively high fees involved. And, of course, courts continued to do a brisk business in the enforcement of the bonds of defaulting collectors.

Not all states adhered to the Ruddock rule. In other states, other considerations seem to have resulted in a far greater role for courts, much earlier. One of the most significant of these was the reluctance of the common law courts to allow too much autonomy to the relatively new creation of state legislatures, the incorporated municipality. Thus, in Bergen v. Clarkson23, the New Jersey court concluded that when a municipal corporation sought to collect its taxes, the summary proceedings available to those jurisdictions acknowledged by common law were not available; the newly chartered cities must sue, and be subject to the oversight of the traditional courts, to collect their taxes unless the legislature had expressly provided otherwise.

The federal government seems not to have embraced the Ruddock approach, although it is difficult to ascertain exactly why. From the first enactment of federal internal taxes, the collector was directed to bring suit in court.24 There is some evidence that this legislation was in fact intended to assure taxpayers that federal collectors would not challenge the validity of the tax itself, since the collector would not be expected to go behind the warrant authorizing his seizure, e.g., Milan Distilling Co. v. Tillson, 17 F. Cas. 280 (N.D. Ill. 1880).

23 6 New Jersey L. 352 (179x). Dillon summarized this approach, which was conveniently consistent with his overall approach toward municipal autonomy: “Municipal power to collect by distress and sale cannot be implied because the state collects its taxes in this manner.” Dillon, Commentaries on the Law of Municipal Corporations, 1st ed. 1872, §656, p. 614. In the earlier, Dillon had acknowledged the Ruddick approach, but qualified it with the notion that “the right to collect by suit should not be taken to be impliedly denied, unless the intention of the legislature, that the special mode prescribed should be the only mode, appears with reasonable certainty.” Id., §653, p. 612. New Jersey apparently joined the ranks of those jurisdictions following the Ruddick rule by providing the power of summary process for taxes in city charters, which was viewed by New Jersey courts as the exclusive remedy. See, e.g., Camden v. Allen, 26 N.J.L. 398 (1857) (on a request for an advisory opinion from the circuit court, holding that an action at law to enforce payment would not be available. In this case, the opinion makes clear that (perhaps counter-intuitively) the debt action would allow a more complete execution against the debtor’s land).

24 Hamilton appears to have preferred direct actions over the traditional methods, because provision therefore was included in several of the drafts of tax legislation proposed for the state of New York attributed to Hamilton, as well as in the federal revenue raising acts passed while he was Secretary of the Treasury. See, e.g., sec. 23 of 1 Stat. 199, 204. Later antebellum federal taxes relied only on the traditional methods of levy, distress and forfeiture, for their enforcement. E.g., Act of July 22, 1813, ch. 16, 3 Stat. 22 (War of 1812 direct taxes), Act of July 24, 1813, ch 24, 3 Stat. 24)(carriage tax).
abuse their power to collect taxes summarily. In his initial proposal for the whiskey tax in 1790, Alexander Hamilton had urged Congress that his proposal included “not only strong safeguards against [revenue officers] being guilty of abuses of authority… [and] even when seizures [which required a warrant from a state or local judge] are made with probable cause… if there be an acquittal of the articles seized a compensation to the proprietors, for injury their property may suffer, and even for its detention, is to be made.” 25 (If these provisions were intended to avoid public criticism, they backfired badly. Although the collectors of the duty on distilled liquor did generally use the federal courts, their use in the early years of the tax required an appearance in Philadelphia—where the only federal court for Pennsylvania was located, by small still owners near Pittsburgh. This aspect of the tax was a major grievance of those protesting in what we have come to call the Whiskey Rebellion.26)

The federal courts, furthermore, seemed to have little difficulty accepting a significant role in the enforcement of customs duties. In the earliest reported case in which the issue was fully discussed in an opinion,27 Justice Joseph Story, acting as circuit court justice for Massachusetts, had little trouble concluding that an action of debt was as appropriate a remedy for the government itself as it was for a private citizen. He acknowledged that such cases in England would have been brought as debt actions only when there was no dispute about the amount owed, and that in cases in which any accounting or valuation was required the suit would be brought by information in the Exchequer, and gave several examples. Rather than fret about whether the jurisdiction of the Exchequer had

25 Richard Hildreth, in his History of the United States of America (first published in New York in 1851, revised in 1879 and reprinted in 1969) points out Hamilton’s purpose: “The plan proposed would be free from the objections principally urged to taxes by excise. No summary jurisdiction would be vested in the excise officers.” p._______. In its report on February 23, 1796, the Committee on Ways and Means resolved that “it would be expedient, after demand made of any tax, (except on goods imported) and a neglect or refusal to pay, to authorize a collection thereof by distress.” No. 91 (4th Cong. 1st Sess.) Finance, Vol. 1 Am. State Papers 385.


27 United States v. Lyman, 26 F. Cas. 1024 (1818) (Case No. 15,647, Circuit Court, D. Massachusetts) (taxes imported at a time when the procedural provisions of the revenue act of 1799, 1 Stat. 627, c.22, controlled)
been given to the federal courts, and if so, what form the suits should take, Justice Story simply observed that they stood for the proposition that import duties were debts of the importer upon importation, and as such debts were clearly actionable. (He simply overlooked the total lack of Exchequer revenue practice in the colonies and states.\[28\] He dismissed out of hand the notion that a bond posted at the time of import (but apparently not paid upon) was the government’s exclusive remedy.

In his holding in Lyman, Story relied on the idea that debts to the government were no less “debts” than private debts.\[29\] It is not clear that Story was correct on this point. The cases he cites in Lyman are all suits on information, brought by the English attorney general in the Court of the Exchequer.\[30\] The possibility that federal courts were not given the power of to hear suits by the federal government against its debtors in the absence of a statute directing the nature of such suits was simply not acknowledged.

Story’s language need not have gone anywhere near so far as it did, because the federal courts were, from the earliest federal revenue acts, given the power to hear disputes arising under them. Indeed, Congress seems from the beginning to have anticipated a greater use of the judicial power of the United States as exercised by Article III courts in administering its revenue laws than most states did.

Story’s stretch to describe a government action to collect a tax as a collectible in a common law action of debt was undoubtedly made to ensure that a jury trial could be had. He did not, however, pause to consider whether the English actions upon which he

\[28\] [Stanley N. Katz, The Politics of Law of America: Controversies over Chancery Courts and Equity Law in the Eighteenth Century, 1971 Perspectives in American History, p. 278, recounting the attempts by a new governor of New York to create an exchequer court in order to secure a juryless tribunal to hear his claim against his predecessor for salary]. The remedies against the King’s debtors in the revenue side of the exchequer included the writ of extent, various aspects of which seem to have no counterpart in practice in any colony or in early federal practice. For brief description of the practice in the Exchequer, see Murray’s Lessee, 59 U.S. at 277-78 and Damsky v. Zavatt, 289 F.2d 46 (2d Cir. 1961).

\[29\] [tendency to misapply cases in which consequences of concluding that tax liability is “debt” for purposes other than collector’s ability to sue, e.g., running of interest, availability of offset, priority in bankruptcy]

relied for the existence of a cause of action at all in the absence of a statute, would have afforded the taxpayer a jury.\textsuperscript{31}

In other cases arising in the early republic, federal courts simply assumed that various types of recoveries related to taxes could be brought as debt actions.\textsuperscript{32} It was never entirely clear, however, whether the taxpayer would be entitled to challenge the assessment. In an action for debt at common law, the amount of the debt must have been certain—suggesting that, if an action in debt would lie, the taxpayer was foreclosed from challenging the amount of, and perhaps in some cases even the fact of, the liability.\textsuperscript{33} Such a position would be consistent with the view that the tax assessment process

\textsuperscript{31} This gap was not fully discussed until the opinion of Judge Friendly in Damsky v. Zavatt, 289 F.2d 46 (2d Cir. 1961). In Damsky, the court granted the taxpayer the mandamus to the district court ordering a jury trial, but limited it to the portion of the suit in the district court relating to a personal judgment against the taxpayer, and not to the portion involving foreclosure of a lien on real estate.

\textsuperscript{32} Jacob v. United States, 13 F.Cas. 267 (1821)(Marshall, as circuit justice, stating that in an action in debt for an unspecified penalty relating to the whiskey tax, a failure to specify the charges in a way that would have been required in an indictment would not defeat a debt action); Meredith v. United States, 38 U.S. 486 (1839)(Story, concluding that the United States could bring a direct action for import duties against importing partners despite the fact that United States had already obtained a judgment on the bonds given for these duties that remained unsatisfied); United States v. Elliot, 25 F. Cas. 1000 (D. Mass.1879)(exploring many aspects of “quasi-criminal” actions for penalties and the effect of Massachusetts law on such proceedings in federal court)

\textsuperscript{33} The law on this point was still unclear as late as 1874, when, in Clinkenbeard v. United States, 88 U.S. 65 (1874), involving a suit on a distiller’s bond under the Civil War distiller’s tax, the taxpayer was allowed to contest the liability. Justice Bradley, for the majority, asserted:

\begin{quote}
It is undoubtedly true that the decisions of an assessor or board of assessors, are of a quasi judicial character, and cannot be questioned collaterally when made within the scope of their jurisdiction. But if they assess persons, property or operations not taxable, such assessment is illegal and cannot form the basis of an action at law for the collection of the tax, however efficacious it may be for the protection of ministerial officers charged with the duty of actual collection by virtue of a regular warrant or authority therefore. When the government elects to resort to the aid of the courts, it must abide by the legality of the tax.
\end{quote}

This view, seemingly uncontroversial to the modern observer, prompted a dissent by four, who believed that the assessment upon which the bond apparently had been based could not be challenged, at least when an appeal to the Secretary of the Treasury had clearly been available:

Such must be the rule, else it will follow that nothing can be collected of the taxpayer in any case where the assessment is for an amount greater than that authorized by law, which is a proposition at war with the whole system of Federal taxation.
rendered its own “judgments” regarding tax liabilities, of equal juridical weight as a court judgment.

The federal courts were able to stick to the position initially outlined by Story. (The relatively rarity of federal internal taxes during the nineteenth century, combined with the general expectation derived from a totally different set of historical traditions that courts would be involved in customs and related admiralty proceedings, probably did much to limit the circumstances in which the courts would see their own political capital at risk.) Other than the initial legislative push given by Congress in the first revenue acts, fed by Joseph Story’s impatience with those who insisted that there was no federal common law for the federal courts to rely upon, there seems to be no obvious explanation for the starkly different positions taken by state courts and by the federal judiciary.34 It did not,

34 Many of the authorities commonly in this line of cases, allowing suits in the name of the United States relate to the collection of penalties and forfeitures related to tax administration, for which judicial oversight of a criminal cause of action would have been required if the courts had balked at entertaining the civil actions. These cases frequently involve difficult (and still not entirely settled) questions about the relationship between civil penalties (for which the cases hold the government may sue for in a common law action in debt) and criminal sanction (which must be brought by information or indictment, and which entail various other procedural complications for the government.) Because these cases for the most part involve statutes which would have required the government to use the courts for criminal process, were the courts inclined to look unfavorably upon the debt action.

The more an action appears to have criminal aspects, the more likely it is that the government has always been required by statute to invoke the power of a court to inflict it[ with the exception of imprisonment for debt]. Thus, in most cases in which the penalty for failure to pay a tax is stated as a direct forfeiture of the object taxed, statutes provided for judicial oversight. In earliest practice, forfeitures could be initiated by government officials or private actors, in suits in the nature of qui tam. Such statutory provisions nevertheless gave rise to questions about exactly who, and in what type of suits could proceed. The very fact that cases can be found challenging the government’s ability to bring such actions suggests how deeply embedded the Ruddock approach became, at least in some jurisdictions. See, e.g., United States v. Chamberlain, 219 U.S. 250 (1911)(allowing an action in debt for penalties under the 1898 Stamp Tax); Stockwell v. United States, 80 U.S. 531 (1871)(holding that the United States could use debt action, rather than a criminal indictment or information, to pursue the double-value forfeiture provided by the act of March 3, 1823, 3 Stat. 781, as specified for all import duties by section 89 of the Act of March 2, 1799, despite the uncertainty in the amount to be collected caused by a scienter requirement related to the double-value forfeiture and despite the additional remedial provisions of the Act of July 18, 1866, 14 Stat. 179); United States v. Colt, 25 F. Cas. 581, 1 Peter’s 145 (D. Penn.1818) (allowing action of debt for unspecified statutory penalty); Jacobs v. United States, 13 F. Cas. 267, 1 Brockenbrough 520 (1821)(allowing suit as action in debt for penalties for violation of provisions of duties on distilling, over defendant’s complaint that only criminal proceeding, with its additional safeguards is appropriate); United States v. Lyman, , United States v. Allen, 24 F. Cas. 772 (1810)(allowing suit in the name of the United States for unspecified import duties to proceed as civil suit under authority of section 89 of the Act of March 2, 1799, 1 Stat. 695, and thus a proceeding by indictment was not required)
however, go entirely unnoticed. In United States v. Tilden,35 Judge, later Justice, Samuel Blatchford asserted that the difference between state practice (not generally allowing suits by collectors) and federal practice, especially after Dollar Savings Bank clearly extended beyond customs matters the practice of allowing suit and treating that suit as independent of any assessment) lay in the nature of the tax:

A scheme of taxation like that found in the federal statutes, where there is imposed by the statute a fixed tax, by a percentage on an amount of money, the elements for ascertaining which are definitely designated in the statute, or a fixed tax of a given amount on a designated object or subject of tax, is a very different scheme from that which prevails generally in the states, where power is confided to public officers to value property, real or personal, and to fix the percentage of tax thereon. There no tax is imposed until the officers act, and no suit for any tax will lie till after such action by the officers.

Blatchford’s explanation, if it rests only on the likelihood that there would be no uncertainty regarding the amount of federal taxes, compared to state taxes, while true to the rhetoric invoked by those whose work he is attempting to reconcile, it is ultimately unsatisfying. But there is a difference among various types of taxes that is relevant. Imposts (that is, customs) require a determination of particular facts that cannot be known until the tax is in fact due, because the tax is generally imposed on the production of an item, or on a transaction involving the item. Property taxes, on the other hand, are imposed on a base that can be assumed to be fairly static, allowing for a determination of the base of the tax, and thus the share of the tax to be paid, long in advance of the determination of “the officers,” in Blatchford’s words, charged with the political task of setting the rate and applying it to the base. It is this task that early courts seem so assiduously to have sought to stay away from.

35 28 F. Cas. 161 (1878) (in an action at law for unpaid income taxes, held that prior invocation of statutory remedies did not preclude action at law: “If the United States [as decided in Dollar Savings Bank] are not to be regarded as bound to resort to the statutory remedy, they cannot be held to be concluded, by a resort to it, from collecting by suit taxes which they have not collected by means of such resort. So, the decision of the assessor … as to the assessment or tax can hardly be denominated a judicial construction, any more than the decisions of the commissioner of internal revenue, which…’can hardly be denominated judicial constructions … [T]he remedy by assessment and collection of taxes, through the machinery of assessors and collectors, is a remedy for the prompt, periodical ascertainment and collection of taxes, subject always to a concurrent right to bring a suit for the tax, such latter right being one which exists both to collect a tax in the absence of any use of the statutory machinery, and to collect it where the statutory machinery has been used and has failed to collect the true amount of the tax.”)
In some sense, notions of separation of powers were never easily resolved when the power to tax was involved. Courts in some instances appear to have feared for their own legitimacy, knowing that if they became too embroiled in the politics of taxation their own credibility would be at risk. In other instances, they simply seem aware that their institutional capacity, and the government as a whole would not withstand the onslaught.36

[...]

[Tax as a legislative function, even for the federal courts:

Oceanic Steam Navigation Co. v. Stranahan, 214 U.S. 320 “in accord with ...settled judicial construction the legislation of Congress from the beginning, not only as to tariff but as to internal revenue, taxation, and other subjects, has proceeded on the conception that it was within the competency of Congress, when legislating as to matters exclusively within its control, to impose appropriate obligations and sanction their enforcement by reasonable money penalties, giving to executive officers the power to enforce such penalties without the necessity of invoking the judicial power.”]

36 Some courts have simply acknowledged that the task of individualized review would be overwhelming: “If left to the judicial department...of the state, to ascertain the amount of indebtedness of the individual taxpayers; and when ascertained, the payment of taxes can only be enforced by a resort to the judicial tribunals of the state, in the same manner as debts are recovered by one citizen from another...the sovereign authority of the state is virtually disrobed of its most important and invaluable rights, of the very essence of sovereignty. The delays and expenses incident to such a system of collecting the public taxes would effectually paralyze the right arm of government, and render it wholly incompetent to the accomplishment of the all important objects for which it was constituted.” E.g., Van de Griff v. Haynie, 28 Ark. 270 (1873)(holding that equalization of property taxes by a board consisting of members of the state senate did not violate the state constitutionally mandated separation of powers, since “[t]he duty of ascertaining taxable values, and of assessing and collecting the taxes thereon, necessarily rests in the discretion of the legislature, and it may perform that duty by its own legislative acts, or throught the agency of such officers or tribunals, as it may appoint for that purpose.”)

Bartlett v. Kane, 57 U.S. (16 How.) 263 (1854) in ruling that if the administrative remedies were not availed of, an importer would get no relief from the courts: “the interference of the courts with the performance of the ordinary duties of the executive departments of the government would be productive of nothing but mischief, and we are satisfied that such a power was never intended to be given to them.”
Taxing Disaster: The Administrative Domain of the IRS

Lily Kahng*

In the aftermath of hurricanes Katrina and Rita, federal, state and local governments will transfer massive amounts to Gulf Coast citizens to aid in relief, recovery, and rebuilding. In addition to addressing the urgent needs for emergency food, shelter, clothing and medical care, governments will provide aid for such diverse purposes as job placement and retraining, business rebuilding, community redevelopment, and prevention against future disasters.

Like death, disaster will not deter the IRS. Prompted by calls from lawmakers, taxpayers, and IRS enforcers, the IRS will gear up its administrative machinery to determine the tax treatment of these government disaster-related transfers. In many cases, whether the payments are exempt from tax or not will turn on the IRS’s interpretation of the general welfare doctrine.

As its name suggests, the general welfare doctrine exempts from income transfers deemed made “for the promotion of general welfare.” The IRS first invoked this doctrine in the mid-1900s to exempt from tax a variety of federal transfer payments, such as Social Security, unemployment, and Medicare. Similar payments by state

* Associate Professor, Seattle University Law School.


2 Some disaster-related payments will be exempt from tax under a new statutory provision enacted after the September 11 terrorist attacks. The new statute explicitly excludes from income only certain types of payments made in connection with certain disasters. As to other government transfers that do not fall within the explicit exclusion, the statute provides that the transfers will be excluded if paid “in order to promote the general welfare.” See I.R.C. 139. Thus, the statute incorporates by reference the general welfare doctrine that is the subject of this Article.

There are many other tax aspects of disaster-related payments that are not the subject of this Article. For a compendium of these, see Francine J. Lipman, Anatomy of a Disaster Under the Internal Revenue Code, 6 FLA. TAX REV. 953 (2005).


governments were also held to be excluded from income by reason of the doctrine.⁷ Since then, the IRS has considered whether the doctrine should apply to a wide variety of government payments whose purposes include witness protection,⁸ adoption assistance,⁹ low-income rental housing assistance,¹⁰ distribution of Indian gaming revenues,¹¹ flood relief,¹² and terrorist attack relief.¹³

The general welfare doctrine is unusual in two important respects. First, as a doctrine of exclusion, it stands out against a backdrop of broad inclusivity. The definition of income is expansive, encompassing “all income from whatever source derived,”¹⁴ all “accessions to wealth, clearly realized, and over which the taxpayers have complete dominion.”¹⁵ Exclusions from income are far and few between, and interpreted stingily by the courts.¹⁶ Nonetheless, the general welfare doctrine paints a broad swath of exclusion across this backdrop of broad inclusivity.

A second, even more extraordinary feature of the doctrine is that it has developed and operated almost entirely outside the formal regulatory, statutory and judicial framework. Virtually all of the legal authority delineating the scope of the doctrine takes the form of IRS guidance of varying degrees of formality--revenue rulings, notices, private letter rulings, general counsel memoranda, technical advice memoranda, chief counsel advices, and field service advices. The general welfare doctrine is the only purely “administrative” doctrine of exclusion that exists in our tax system.

As developed through the administrative process, the contours of the general welfare doctrine are opaque at best, and probably more fairly characterized as incomprehensible and shifting. The IRS appears to articulate several underlying principles: The payments must be made by a governmental unit for the promotion of

---

⁸ See IRS Field Service Advice CC:TL-N-2928-92, 1992 FSA LEXIS 70.
⁹ See IRS Chief Counsel Advice 200021036, 2000 IRS CCA LEXIS 31.
¹⁰ See, e.g., IRS Chief Counsel Advice 199948040, 1999 IRS CCA LEXIS 246.
¹¹ See IRS Technical Advice Memorandum, PLR 9717007, 1997 PLR LEXIS 81.
¹⁴ I.R.C. Section 61(a).
¹⁶ See Commissioner v. Schleier, 515 U.S. 323, 328 (1995) (“exclusions from income must be narrowly construed”); United States v. Wells Fargo Bank, 485 U.S. 351, 354 (“exemptions from tax are not to be implied; they must be unambiguously proved”).
general welfare, health or safety. They must generally be made on the basis of a determination of financial need by the recipient. They cannot be compensation for the performance of services. They cannot provide support for businesses.

At the same time that it has embraced these principles, however, the IRS has in some cases departed from them, often with no explanation for the departure. Thus, for example, though need-based governmental payments are usually excluded from income, the IRS has taxed such payments received from foreign governments. Though payments are usually required to be made on the basis of financial need in order to be exempt, the IRS has excluded non-needs-based payments to blind individuals or individuals adopting children. Though the exclusion is usually predicated on the lack of a compensatory quid pro quo, the IRS has excluded welfare payments requiring the performance of work by recipients. Despite the general requirement that excludable payments cannot provide business support, the IRS has excluded small business grants paid by an Indian tribe to its members.

This Article will be the first to undertake a comprehensive examination of the general welfare doctrine. Obviously, one purpose of this analysis will be to gain a better understanding of the contours of the doctrine and its theoretical underpinnings. More importantly, however, the Article will identify the ways in which the doctrine has been applied inconsistently or irrationally, and explore the possible causes of, or biases

---


18 See, e.g., Rev. Rul. 78-170, 1978-1 C.B. 24 (Ohio state payments to low-income elderly or disabled residents to reduce winter heating costs are not includible in income). Compare Rev. Rul. 76-131, 1976-1 C.B. 16 (Alaska state payments to elderly residents continually residing in state for 25 years, without regard to financial status, are includible in income).

19 See, e.g, IRS Chief Counsel Advice 200227003, 2002 IRS CCA LEXIS 45 (property tax abatement given to elderly citizens in exchange for performance of volunteer services is includible in income).


22 See Rev. Rul. 57-102, 1957-1 C.B. 26 (non-needs-based payments to blind individuals); IRS Chief Counsel Advice 200021036, 2000 IRS CCA LEXIS 31 (non-needs-based adoption assistance payments).

23 See Notice 99-3, 1999-1 C.B. 271 (payments under Temporary Assistance to Needy Families, which requires the performance of work, excluded).


25 The only article to address the doctrine in some depth was written 30 years ago. See Charlotte Crane, Matching and the Income Tax Base: The Special Case of Tax Exempt Income, 5 AM. J. TAX POL’Y 191, 233-36 (1986). Moreover, the article’s analysis of the doctrine is incidental to the article’s main thesis, which focuses on coordinating the tax treatment of payors and payees of exempt income. See id.
contributing to, these disparate applications. Economic class is an explicit factor in the doctrine’s application, but the Article will explore whether non-explicit considerations of class, as well as race and gender, may also play a role.

The Article will also examine the different administrative processes by which different types of IRS guidance are produced, also with a view toward better understanding and explaining outcomes under the doctrine. This in turn will lead to observations about the ways in which informal administrative practices can affect substantive outcomes, and the difficulties of achieving greater consistency in these practices. The Article will conclude with recommendations for greater transparency and fairness in these practices, along with directions for further research.
“What is Fiscal Responsibility? Long-term Deficits, Generational Accounting, and Capital Budgeting”

[Note to Participants in the 2006 Critical Tax Conference:

[I presented this working draft at the NYU Tax Policy Colloquium two years ago. After receiving feedback on this draft, I decided to split the paper into four separate projects: an article critiquing so-called generational accounting, an article advocating the adoption of capital budgeting by the federal government, an article on sustainable fiscal policy, and an article on current generations’ responsibilities to future generations.

[I completed the first article last year (publishing it at 58 Tax L. Rev. 275). I am now turning my attention to writing the article on capital budgeting. The new article’s tentative title is “Spending Responsibly in Times of Crisis (and in Normal Times, too): Creating a System of Capital Accounts for the Federal Government.”

[For the purposes of this conference, therefore, I would appreciate your feedback specifically on how to develop the ideas in Section III of this draft, which presents the basic concept of capital budgeting and some of its strengths and limitations. For background, you will probably find it useful also to read the Introduction and Section I. You can skip or skim Section II. (If you happen to look at the very short fourth section and have some thoughts, I’d be grateful for any feedback before or after the panel, as that section is the starting point for the fourth of the papers described above.)

[Thank you.]

Neil H. Buchanan*
Center for Law & Justice
123 Washington Street
Newark, New Jersey
neilb@andromeda.rutgers.edu

* Assistant Professor of Law, Rutgers School of Law – Newark. J.D. University of Michigan Law School, Ph.D. in Economics, Harvard University. I would like to thank Reuven Avi-Yonah, Michael C. Dorf, James R. Hines, Jr., Kyle D. Logue, and Daniel N. Shaviro for helpful comments and discussions.
Table of Contents

Introduction......................................................................................................................... 1

I. Traditional Fiscal Deficit Measures............................................................................. 8
   A. Preliminary Matters................................................................................................. 9
      1. Debt and Deficits ............................................................................................. 10
      2. Dollars vs. Percentages ................................................................................... 11
      3. Federal vs. State and Local ............................................................................. 12
   B. Cash-Flow Deficit Measures.................................................................................. 12
      1. On-Budget and Off-Budget ............................................................................. 13
      2. Cyclical Adjustment ......................................................................................... 14
      3. Unfunded Liabilities ......................................................................................... 16
      4. Extended Budget Projections ........................................................................... 17
   C. Effects on Consumption and Investment ............................................................... 17

II. Fiscal Gaps and Generational Accounting................................................................. 19
   A. Computing the Accounts....................................................................................... 21
      1. The Fiscal Gap .................................................................................................. 21
      2. Lifetime Net Tax Rates .................................................................................... 24
   B. Weaknesses in Generational Accounting............................................................. 25
      1. Paying Down the Debt ..................................................................................... 25
      2. Benefits Not Counted in a Generation’s Accounts .......................................... 27
      3. Demographic Trends ......................................................................................... 28
   C. The Policy Imperatives ........................................................................................ 30
      1. The Norm of Generational Equality ............................................................... 30
      2. The Norm of Shared Sacrifice ......................................................................... 30
   D. Generational Accounting Without Generations? .................................................. 31
      1. Another Baseline Problem .............................................................................. 31
      2. The Fiscal Gap and the Annual Deficit ............................................................ 35
      3. Making the Fiscal Gap Disappear ................................................................... 36

(Table of Contents continued on next page)
Table of Contents (continued)

III. Capital Budgeting ........................................................................................................ 38
    A. Capital Budgeting ................................................................................................. 38
        1. Definitions ....................................................................................................... 38
        2. Using Capital Accounts ............................................................................... 40
    B. Beyond Capital Budgeting ................................................................................ 42
    C. Back to Small Decisions .................................................................................... 43

IV. What Does it Mean to Be Fair to Future Generations? ........................................ 43
    A. How Much is Enough? ....................................................................................... 44
    B. The Real Inter-Generational Issues .................................................................. 45

V. Conclusion ................................................................................................................. 46
Introduction

The decisions that we make today regarding taxes and government spending have profound effects not only on those of us living today, but on future generations as well. These effects arise for two reasons. First, our current taxing and spending decisions help to determine how the economy’s productive resources will be used now and in the future—whether, for example, a piece of land becomes the site for a casino, a day care center, or a cancer research institute. Second, the laws that we pass generally commit the government to courses of action that can last well into the future. While it is surely true that some laws that claim to set policy for years in advance do not really do so (such as tax policies passed in 2001 that purport to expire in 2010), others surely represent commitments from which the government would have some difficulty withdrawing.

These effects of taxing and spending policies—on the current uses of productive resources, as well as on the somewhat-credible commitments to future policies that they frequently represent—ought to be of concern to anyone whose time horizon extends past the current fiscal year. For those who care about the state of the world that we leave to future generations, the challenge is to find an analytical framework with which to predict the impact of current policy choices on future standards of living. Typically, these analytical frameworks are found in macroeconomic analysis of fiscal policy.

Macroeconomics does not, however, figure prominently in legal analysis. While the legal literature that relies on microeconomics has mushroomed over the last two
decades or so, macroeconomics has largely remained in the background. With the occasional exception, legal scholars have tended not to include in their analyses the aggregate economy within which microeconomic efficiency analysis operates. Certainly, some very good work has been done analyzing proposed constitutional amendments to constrain government spending or to limit tax increases; but these tend to be self-contained analyses—and they need not challenge the orthodox view that annual budget balancing is the essence of fiscal probity. 

In public policy debates, of course, most analysts will acknowledge that some solutions are not “realistic” because they are too expensive. But what does it mean to be too expensive? That question, which lies at the heart of fiscal macroeconomics, is considerably more complex than it might first appear. Some items with very large price tags are not considered to be too expensive under certain circumstances. For some parents, a tuition bill exceeding $30,000 a year—while a source of frustration, to be sure—is not too expensive to send their bundle of joy to a high quality college or university.

When discussing public spending, however, the notion of “too expensive” can often take on a simple-minded tone. If it “increases the deficit,” it is often thought to be per se too expensive. Since every government program, by definition, increases the deficit above where it would otherwise be, only literally self-financing programs can pass muster under such a barren decision-making regime. For legal analysts, therefore, any

1 Mark R. Kelman, Can Lawyers Save the Economy? Stan. L. Rev., 1993 (surveying recent macroeconomic literature, noting that many current models blame poor macroeconomic performance on incomplete contracting, and suggesting that lawyers can help to solve such problems)
2 See, for example, Nancy C. Staudt, Constitutional Politics and Balanced Budgets, 1998 U. Ill. L. Rev. 1105 (1998) (arguing that it is important to balance the budget, but a constitutional amendment is unnecessary because the budget balancers have won the political debate).
discussion of proposed changes in policies would have to proceed either from an agnostic fiscal viewpoint—“Whether this policy would still be desirable, after taking the deficit into account, is beyond the scope of this analysis”—or must explicitly make the claim that the proponent’s policy is “not too expensive.”

The latter argument appeared to become somewhat easier to make when federal deficits turned into surpluses in the late 1990s, and again when the aftermath of September 11, 2001 seemed to suspend all previous rules on spending. The basic arguments against deficit spending are unlikely ever to go away, though, because of the widely held belief that fiscal responsibility is synonymous with annually balanced budgets (or, more extremely, with zero public debt). Indeed, even in the face of a weakening economy and the threat of a double-dip recession, President George W. Bush decided to show his economic seriousness during his August 2002 economic summit in Waco, Texas by announcing that he would refuse to spend $5.1 billion that Congress had already approved for domestic security and the military. Explaining that decision, he declared: “More money spent in Washington means less money in the hands of American families and entrepreneurs, less money in the hands of risk takers and job creators.”

Responding to the political focus on annual cash-flow deficits (and their many variations), in the early 1990’s the economist Laurence Kotlikoff and his frequent co-authors Alan J. Auerbach and Jagadeesh Gokhale developed a theory that was motivated

---

3 See, for example, Richard W. Stevenson, Budget Deficit Is Said to Be $159 Billion, N.Y. Times, October 25, 2002, at A27 (“The return to red ink brought an end to the four-year period in which surpluses and the promise of more had left both parties almost giddy with the possibility of addressing the nation’s needs without painful tradeoffs.”)

4 Note that $5.1 billion is approximately one-fourth of one percent of the overall federal budget and less than 0.05% of annual U.S. GDP.

by understandable concerns about the long-term effects of federal fiscal commitments. Dubbed Generational Accounting,\(^6\) the resulting model—whatever its merits and shortcomings (discussed in some detail below)—would put an even greater constraint on legal analysts or anyone else who might propose a change in public policy. No longer would it be enough to prove that a policy would not increase the current deficit; Kotlikoff et al. would offer opponents of government activism an even more powerful trump card. Generational Accounting (hereinafter GA) requires not merely that there be enough money to pay for the program today, but it also requires that the program be “affordable” into the indefinite future, as measured by long-range budget forecasts. Adding a political (and somewhat emotional) slant to the discussion, GA’s proponents further suggest that anything that is not affordable in this way is a transfer to the (voting) living generations from (politically defenseless) unborn generations.

The basic notion behind generational accounting is quite simple—though such conceptual simplicity makes the intractable measurement problems discussed below all the more frustrating. If current spending and tax laws were to stay in effect in perpetuity, the flows of expenditures and revenues would vary depending on future economic growth, population changes, weather patterns, medical developments, etc. GA makes some very basic assumptions about the directions and magnitudes of the most important

of those future trends, applies the accounting concept of net present value to bring all of those projected future expenditures and revenues into current dollar terms, and then computes the different lifetime net tax rates implied by those calculations for different generations. Hence, Kotlikoff and his co-authors claim that GA measures whether current generations are being “fair” to those that will follow.

*Tax Notes*—the periodical of record for tax practitioners, policymakers, and academics—has carried at least eight articles that in some way deal with GA since 1991, when the theory was first introduced. To date, the appearances of GA in the legal literature has not been extensive, but this may be changing. Professor Daniel Shaviro of

---


New York University Law School has written, in addition to a short article in *Tax Notes*, two books (and has finished a third) that to some degree would import GA into legal analyses of Social Security policy and Medicare policy—and possibly into all legal analysis of fiscal policy. Suggesting a long-term fiscal imbalance that makes “the current policy fiscally unsustainable,” Shaviro asserts that it is necessary to think about these issues through the lens of GA—and such an analysis, we learn, shows a grim future indeed, as described below.

The stakes in this debate, therefore, could hardly be higher. Anyone who wishes to write about Medicare or Social Security, and indeed anyone who might ever wish to suggest that the government should spend money on any project, might plausibly be forced to contend with the implications of GA.

If GA lived up to its billing, of course, it would simply be good policy to use it as a starting point for fiscal analysis. Unfortunately, GA is not a neutral analytical tool that is used dispassionately to assess the fiscal consequences of a government project. Instead, GA is based on highly contestable assumptions, arbitrary analytical choices, and manipulable policy projections that fatally compromise its analytical value.


While some of these references are merely *pro forma* (e.g., McCaffery, Garrett, Newman), and Kornhauser directly criticizes the theory, several of the articles (esp. Brody’s pieces and Forman’s article) actively endorse GA. 9 *Supra* note 7.


12 Shaviro, *supra* note 7, at 715.

13 *Id.* at 716 (“Perhaps the best tool for enhancing our understanding of who wins and loses from alternative reforms is generational accounting….”) Shaviro is not, however, explicitly wedded to GA to the exclusion of all else. Instead, he has suggested that we should use GA as well as other theories to learn as much as we can from a variety of approaches.
Understanding these limitations is essential for anyone who wishes to be able to respond to the suggestion that their favored policy is, in this much broader sense, “too expensive.”

The debate, therefore, is not over whether the future matters—that is, this is not a debate between the grasshoppers and the ants. Clearly, we must always think carefully about the future consequences of our fiscal policies. The question is how to think about the future—what we would like to bequeath to future generations and how best to deliver it.

This article assesses three basic approaches to assessing the future effects of the government’s fiscal policies: traditional measures of the deficit, measures associated with Generational Accounting, and measures derived from applying Capital Budgeting to the federal accounts. I conclude that Capital Budgeting is the best of the three approaches and that Generational Accounting is the least helpful. Acknowledging that there might be some value in learning what we can from a variety of approaches to analyzing fiscal policy, I nevertheless conclude that Generational Accounting is actually a misleading or—at best—empty measure of future fiscal developments. The best approach to providing for the future is thus to apply careful cost-benefit analysis through old-fashioned Capital Budgeting to our spending and taxing decisions; but if political pressures prevent the adoption of a federal capital budget, we would be best served by continuing to use our current deficit measures, with some minor adjustments.

Because we are attempting to peer into the future, any measure of the effects of fiscal policy will be imperfect. Choosing among those imperfect alternatives is the focus of this essay.
I. Traditional Fiscal Deficit Measures

Even the most casual observer of U.S. political debates cannot have missed the fact that our politicians are obsessed with “the deficit.” After decades in which Republicans regularly attacked Democrats for their spendthrift ways, Democrats delighted in turning the tables in the 1980s, as Ronald Reagan presided over the largest nominal peacetime deficits in American history. Undaunted, conservative Republicans insisted that they were the truly responsible fiscal custodians, culminating in the promise to balance the budget in 1994’s Contract With America.

Capitulating to the political heat generated by this headline-grabbing issue, former President Clinton agreed in 1995 that he, too, was committed to balancing the budget.\textsuperscript{14} When the budget moved from deficit, to balance, and then to surplus under his watch, Clinton never missed an opportunity to take credit for this “achievement.” With their own party’s leadership having abandoned them, even the most progressive members of the Democratic Party became committed budget balancers. Senator Russell Feingold, for example, eagerly points out that he opposed President Clinton’s proposed middle-class tax cut in the early 1990’s. “I was for deficit reduction.”\textsuperscript{15}

The politics of fiscal deficits can, therefore, change rapidly. In an early draft of this article, written in the late autumn of 2002, it was plausible to argue: “Officially,

\textsuperscript{14} Indeed, Clinton’s pre-1994 actions showed that he was strongly predisposed to the balanced-budget mantra, as he immediately jettisoned his post-election plans for long-term capital spending and instead pushed through a major tax increase. (For those who approve of tax progressivity, however, one can at least note that Clinton’s tax bill was top-loaded.)

\textsuperscript{15} Matthew Rothschild, \textit{The Progressive Interview: Russ Feingold}, The Progressive, May 2002, at 31. Feingold’s reasoning is, at best, puzzling: “I think it’s a mistake from the point of view of our economy and also as far as gaining credibility with the American people if we don’t try to avoid deficits. Whatever organization you are with, whether it’s an environmental group, right-to-life group, Communist Party, you all are going to have to pay the bills. You establish credibility with people when you show them that whatever your ideology you will take care of their dollars in a businesslike way.” \textit{Id.} Feingold thus confuses (among other things) government budgeting with personal budgeting.
therefore, there are no longer any major political voices in American politics that argue in favor of deficit spending. At that point, as the fiscal ink turned red again, a new consensus arose that the return to short-run deficits was not a serious problem—but that long term deficits represent a virtual “cancer” eating away at future prosperity. With the onset of the 2004 presidential campaign, of course, finger-pointing about the deficit has risen to a fevered pitch.

To some degree, this political obsession with annual deficits is quite surprising, because the economic arguments in favor of deficit spending in various circumstances are well known and, to a large extent, uncontroversial. While it is certainly possible to argue that, in spite of the economic case in favor of deficits in some circumstances, there is a stronger political case against them, it is at least worth remembering what the economic issues are.

A. Preliminary Matters

Economists differentiate between two types of variables: stock variables and flow variables. The difference between the two has to do with the passage of time. Stock variables can be measured at a moment in time, while flow variables are only meaningful

16 Those who do oppose the orthodoxy are not necessarily at the liberal end of the spectrum. Former Congressman Jack Kemp was once the most prominent political advocate of so-called Supply Side Economics, which holds that low tax rates are much more important than balanced budgets in generating high economic growth rates. [cite]
17 Neil H. Buchanan, Providing for Future Generations: Generational Accounting, Capital Budgeting, and Budget Deficits, unpublished manuscript (on file with author), November 25, 2002, at 8 (footnote in original). Compare also Prof. Staudt’s conclusion in 1998 that budget balancers had prevailed in the political debate. See note 2, supra.
19 Alan S. Blinder, Is the National Debt Really—I Mean, Really—a Burden? in James S. Rock, ed., Debt and the Twin Deficits Debate, 1991, at 209-25 (arguing that, even though the deficits of the 1980’s and early 1990’s were relatively minor, there could be no political peace unless everyone agreed to balance the budget).
Thus, distance is a stock variable, while speed (miles per hour, for example) is a flow variable. In economics, common stock/flow distinctions include prices (stock) versus inflation (flow), and wealth (stock) versus income (flow).

1. Debt and Deficits

In government accounting, debt is a stock variable, because it measures the total amount of money at any given moment that a government owes its creditors. The deficit is a flow variable, measuring the net amount of new borrowing that the government has engaged in during the course of a year (or any other unit of time). Deficits are, properly measured, the change in debt as time passes.

Importantly, the total amount of federal government debt that exists at a given moment is also tautologically equal to the total number of dollars of Treasury securities in circulation at any moment. Since the federal government borrows money by selling Treasury Bills, Notes, and Bonds, the face value of those outstanding securities is equal to the National Debt.

Here, however, is one of the first places where reality and theory diverge. The different agencies of the federal government often hold each other’s debt instruments, so the net federal debt is lower than the number of bonds that have been issued but not redeemed. The difference is not trivial. While the infamous “National Debt Clock” (the ever-rising digital readout of the nation’s supposed indebtedness, expressed both in the aggregate and as “Your Family’s Share”) shows an outstanding federal debt of over $7.1 trillion in April 2004, just over $4 trillion of that total was held outside of the federal

government’s own offices.21 This was why the National Debt Clock did not initially decline when the federal government ran surpluses in the late twentieth century. Even though the government was extinguishing debt held by the public, the total number of bonds in existence was not going down, so “the national debt”—by that meaningless measure—was not shrinking.

2. Dollars vs. Percentages

The common practice of expressing debt and deficits in total dollars rather than as percentages of national income can also be highly misleading. Politicians in the late 1980s talked of “$200 billion deficits as far as the eye can see” as if that was an unimaginable calamity. In fact, given that nominal GDP doubles roughly every twelve years, $200 billion annual deficits would be trivial in relatively short order.

In the U.S., deficits as a percentage of GDP peaked in the mid-1980s at roughly 6%, and the publicly-held debt peaked at around the same time at 60% of GDP. This was, of course, only the recent peak and was not even close to the 125% debt level reached at the end of World War II, when we wisely spent enormous sums of borrowed money to finance the war against the Axis powers. This ratio had steadily fallen to the point where it was below 50% by the early 1980’s, and then rose for over a decade before falling again in the late 1990’s and early in the new century. The current return to deficit spending finds projected deficits at about 4.2% of GDP in 2004 ($477 billion), and the federal debt at 38.2% ($4,385 billion).22


22 Congressional Budget Office, CBO’s Current Budget Projections (March Baseline Projections), March 2004.
3. Federal vs. State and Local

A third measurement issue carries more direct implications for policy debates. Typically, commentators separate the federal debt and deficit from the state and local fiscal positions. Since the state and local sector tended until very recently to run aggregate operating surpluses, the decision to exclude the state and local sector when discussing the “government” deficit and debt naturally made the situation look worse—while the current situation of chronic state deficits is ignored by federal measures of borrowing. Foreign economists view this practice as nothing less than bizarre, because the macroeconomic consequences of debt and deficits surely do not depend on the hierarchical level of the government entity that is doing the borrowing.23

In addition to being logically incoherent, this practice has perverse policy affects as well. When national politicians view their job as reducing the federal deficit or debt, they are tempted either to ignore the consequences of their decisions on lower levels of government or even deliberately to shift spending requirements downward.24

B. Cash-Flow Deficit Measures

While the issues discussed briefly above have important implications for discussing the status of fiscal policy in the United States, the discussion that follows attempts to follow the current norms in describing deficit measures. Even within the

---

23 Wynne Godley, Seven Unsustainable Processes: Medium-Term Prospects and Policies for the United States and the World. Special Report: The Levy Economics Institute (revised Oct. 5, 2000), at 2 (“In the United States the public discussion of fiscal policy concentrates almost exclusively on the operations of the federal government. Yet state and local governments account for about a third of all public expenditure and taxes; moreover, their budgets are generally in surplus so that these authorities are now in substantial credit . . . . In what follows, government inflows and outflows--and debts--will always refer to the operations of the "general government" (the combined federal, state, and local governments).”)

federal-only measures, however, there are significant disagreements about what is the true measure of fiscal policy.

1. On-Budget and Off-Budget

Even if one looks only at the federal government, the annual deficit is more manipulable than it might seem. A spending program can exist in a netherworld outside of the official budget simply by act of Congress. There need be no economic rationale for the decision. The 1991 Gulf War was carried off budget, for example, and the current operations in Iraq are being funded by emergency appropriations. By far the biggest off-budget item, of course, is the Social Security system. The current surplus in that system (approximately $161 billion in 2004, or 1.4% of GDP) makes the total deficit smaller than the on-budget deficit ($638 billion, or 5.6% of GDP), but when Social Security starts to run deficits after the next decade or so, the on-budget deficit will be smaller than the total deficit.25

The debate about whether the Social Security Trust Fund has any meaning is, of course, an important factor in determining whether the on-budget or total deficit is the proper measure. Because I conclude that the total deficit is the better of the two, in that it measures the amount of money that the federal government is draining from the financial markets in a given year, I will focus on that measure of the deficit and possible ways to improve it.26

25 All estimates in this paragraph are from Congressional Budget Office, CBO’s Current Budget Projections (March Baseline Projections), March 2004.
26 For a persuasive argument that the Social Security system should not be seen as an individualized benefit plan but rather as a redistributive fiscal program, see, Deborah A. Geier, Integrating the Federal Tax Burden on Labor Income, 98 Tax Notes 563, 574 (Jan. 27, 2003) (citing generally Patricia E. Dilley, Taking Public Rights Private: The Rhetoric and Reality of Social Security Privatization. 41 B.C. L. Rev. 975 (2000)).
2. Cyclical Adjustment

For macroeconomists, one of the most important measurement issues in deficit accounting is adjustment of the deficit for changes in the health of the economy. When the business cycle turns, tax receipts and government expenditures naturally change along with the GDP. Recessions bring lower revenues and higher expenditures, and boom times do the opposite. When comparing deficits at two points in time, therefore, it is important to ask, “What would the deficit be today if the economy were fully healthy?”

What is means to be “fully healthy” is, of course, a matter of contention. Nevertheless, there is a widely-accepted measure of the cyclically-adjusted deficit known as the Standardized-Budget Deficit, computed by determining the flows of revenues and expenditures if the unemployment rate were at its trend rate. In 2003, because of the lingering effects of the recession, the unadjusted total deficit was $375 billion (3.4% of GDP), while the Standardized-Budget Deficit was $313 billion (2.8% of GDP).27

Failing to adjust the deficit for cyclical effects leads to two problems. First, it confuses cause and effect. Improvements in the economy cause decreases in the non-adjusted deficit; but decreases in the cyclically-adjusted deficit (all else constant) cause the economy to decline. Second, and far more importantly, it causes perverse policy moves, as a worsening economy causes the deficit to rise, and attempts to reduce the deficit with further cuts in spending (and perhaps increases in taxes) will only further weaken the economy. President Bush’s symbolic refusal to spend money that Congress had allocated, noted above, clearly demonstrates this perversity.28

28 See notes 3-5, supra, and associated text.
Nor is this failure to comprehend simple macroeconomics confined to the United States. The Japanese economy went into its first of several severe downturns in 1989. By 1996, with the domestic economy still in deep trouble, Japanese policymakers relied on more budget-cutting and tax increases to improve the economy—the macroeconomic equivalent of “bleeding” a patient to restore her to health. Yet policymakers there and elsewhere remained puzzled by their patients’ continued ill health.29 “[M]any economists believe that Japan’s long stagnation in the 1990s largely reflected timid policymakers unwilling to boldly use the levers of fiscal and monetary policy.”30

The practical consequences of failing to adjust for the business cycle are especially severe for state and local governments, most of which operate under (modified) balanced budget requirements.31 When the economy is strong (which means that, by definition, workers are scarce because their prospects are so good in the private sector), states flush with money compete with prosperous companies for workers and other economic resources. Roads are torn up and re-built precisely when the disruption from such projects is the most damaging, as the overburdened highways are filled with vehicles carrying the evidence of economic prosperity. Then, when the economy

29 “Exasperation as Tight Budgets Don't Deliver Growth...” Nomura Securities Research Report, August, 1996 (“Low inflation and trimming of fiscal deficits have always been regarded as a foolproof recipe for economic growth. However, that conventional wisdom has been turned upside down in the past few years, as politicians in developed economies have grown exasperated by the failure of high growth to materialize despite their belt tightening efforts. Average real GDP growth in the major industrialized nations was 4.0-4.5% in the 1970s, and 3.0-3.5% in the 1980s- but in the 1990s many believe the figure will be a meager 2.0%.”)


31 See discussion in “Capital Budgeting,” below.
weakens, states see their tax revenues decline, lay off workers, and leave highways in
disrepair. It would be difficult to design a more perverse system.  

### 3. Unfunded Liabilities

The closest one comes in the traditional deficit debates to the central issues of
long-term budgeting (which are the focus of Generational Accounting) is the discussion
of “unfunded liabilities.” The basic idea is that government projects that involve
spending in the future can be thought of as liabilities that must be accounted for when
looking forward. This sensible observation, though, can only be useful inasmuch as
projects have dedicated financing. If a high school is built with proceeds from a bond
sale, for example, the liability is funded if the school district commits the funds in future
budgets to cover the bond payments. Otherwise, the project is unfunded.

Since most government programs are not financed through dedicated funds, of
course, virtually any project could be called an unfunded liability. The Interior
Department, the Army Corps of Engineers, etc., are all unfunded, and they will continue
to be so for as long as they last. Hence, as appealing as the idea is, making the unfunded
liabilities concept operational is daunting at best. Estimates of unfunded liabilities are
also highly responsive to changes in the law, and their size can dwarf the rest of the
budget. The estimate of unfunded liabilities in the Social Security system after the
change in withholding taxes in the early 1980’s, for example, swung from several trillion
dollars in unfunded liability to several trillion dollars in unfunded surplus. Generational
Accounting is arguably an improvement on the arbitrary nature of these estimates, as I

---

32 This is not to say that highways are the be-all and end-all of economic spending. In this analysis, they are simply the most intuitive example of public spending on infrastructure.
33 Social Security Trustees, cited in Buchanan, infra note __.
discuss below; but GA’s other shortcomings ultimately make it an unappealing alternative.

4. **Extended Budget Projections**

All of the budget measures discussed above are measured in annual terms. It is possible, of course, to use a different arbitrary time period in such an analysis. During the Clinton Administration, it became common to provide ten-year projections of budgets, to allow policymakers to look into the relatively foreseeable future and determine whether a budget or tax measure was likely to become more or less manageable over time. The current Bush Administration has discontinued the ten-year projections in favor of five-year projections, a move that has generated suspicion that the full costs of their policy proposals are “back-loaded.” It is, of course, possible to back-load even on a ten-year horizon.

The arbitrary nature of these cutoffs is, as discussed in the next section, a strong argument in favor of adopting an infinite-horizon model along the lines of Generational Accounting. Nevertheless, because of the critical problems in lengthening the time horizon, the ten-year projections are probably the best compromise available. This point will be taken up below.

C. **Effects on Consumption and Investment**

The payoff for making these adjustments to the measurement of the fiscal deficit comes in analyzing the effects of current deficits on the use of society’s productive resources (labor, machinery, factories, land, etc.). If the government hires resources to build or produce goods and services, and if those resources would have been used to
produce something of value to private citizens, then the government has “crowded out” private activity. If the government crowds out private consumption, then there is at worst no effect on future generations, because private consumption would not have benefited future generations in any case. In such a situation, the government can make future generations better off if it replaces private consumption with public investment, or it can simply substitute one kind of consumption for another—leaving future generations unaffected.

The serious concern, of course, is that the government will not crowd out private consumption but will, instead, crowd out private investment. If the government wastes money that would have gone toward private investment in productive equipment, for example, then the future standard of living of the country is compromised. In fact, even if the government invests resources in productive assets, it can still make matters worse if the assets it crowds out would have been more productive than the government’s investment projects. What is often forgotten, however, is the other possibility—that the government might crowd out a private investment project with a public investment project that is even more productive. Rather than building a strip mall, for example, resources might be used to build a children’s hospital. 34

While there is no precise way to know the exact nature of these tradeoffs, the fundamental question could not be more important: What effect will the government’s decisions today have on the economy’s productive capacity—and thus the real standard of living—tomorrow? It is here that our concern for future generations should be

34 Although the discussion here focuses on spending projects, taxing decisions can be analyzed in precisely the same fashion. Every aspect of the tax code has the potential to change consumption decisions into investment decisions, investments into consumption, less productive investments into more productive investments, etc.
concentrated. Using annual accounts (even properly measured accounts), however, will still arguably not capture the full effects of our taxing and spending decisions.

In summary, if we are to improve public discussion of the government’s fiscal situation, the least radical alternative would be to adopt a cyclically-adjusted deficit (preferably for the entire government sector, though that appears unlikely). Such a measure, while still imperfect, would allow policymakers to focus on the possible crowding out caused by its annual fiscal policies. Longer-term forecasts would still be valuable in many instances. While imperfect, such measures provide policymakers with useful guidance in looking past the current fiscal year.

II. Fiscal Gaps and Generational Accounting

The structure of Generational Accounting is built upon some very appealing foundational arguments. First, as noted above, deficit accounting is arbitrary, because there is nothing special about a year as the unit of analysis. Indeed, arbitrarily aiming to balance annual books can introduce its own set of bizarre games. After taking office in 2001, the second Bush administration changed the national accounts such that some corporate tax revenues would be credited in October 2001 rather than in September of that year. Because the federal government’s fiscal year runs from October 1 through September 30, this move reduced the (then-projected) surplus for fiscal 2001 and increased it for fiscal 2002.

35 Some states use biennial budgeting, which is no more nor less arbitrary than annual budgeting.
This gamesmanship came to light after the 2001 tax cut was passed, when it appeared that the on-budget surplus might actually slip into deficit for the first time in several years. Democrats were quick to accuse the new administration of fiscal irresponsibility, and the administration quickly assured everyone that the on-budget surplus would still be $8 billion.\(^\text{36}\)

Second, even without any such political games, there is no good analytical reason to assess government programs on an annual basis. Projects that last longer than a year should be analyzed in their relevant time frame. Surely, a year is far too short a time in which to measure meaningfully the impact on the economy of the vast majority of government programs and tax policies.

Of course, once one realizes that a year is arbitrary, one must also recognize that there is no non-arbitrary alternative. The infinite future is out there, and perhaps the best way to proceed is to use the simple financial concept of net present value discounting to bring all future receipts and expenditures into one current estimate. This approach has an added benefit, in that it avoids the issue of whether a program has a dedicated financing mechanism. While it arguably makes sense to compare the long-term planned expenditures and expected receipts for something like the highway trust fund, the majority of government programs can only meaningfully be assessed in the aggregate, because the FBI (for example) is not supported by its own tax regime.

---

\(^\text{36}\) In a $10 trillion economy with a $2 trillion budget, the idea that we can predict an $8 billion surplus—less than one half of one percent of spending—stretches credulity. Such a small number is little more than rounding error.
A. Computing the Accounts

1. The Fiscal Gap

The fundamental analytical achievement of the generational accounting framework is its attempt to compute an aggregate, discounted federal\(^{37}\) deficit or surplus into the infinite future. Estimates generated using a GA framework have been included in federal budget documents,\(^{38}\) and Professor Kotlikoff has provided a great deal of input to the work of the Congressional Budget Office.

The basic logic of GA is quite simple. Imagine that current law remains unchanged indefinitely. What are the likely paths of government spending and tax receipts, given expected trends in population, economic growth, etc.? Given those likely paths, what is the aggregate gap between spending and tax revenues into the infinite future? Taking the analysis one step further, it is then theoretically possible to estimate the net amount of money that an individual will pay in taxes to their government over their lifetime.

The appeal of moving to long-term budget calculations loses its luster rather quickly, however, in the face of the complexity of long-term budget estimation. Consider the inputs necessary to generate a GA estimate. Tax receipts for each future year must be calculated on the basis of estimates of the number of taxpayers, their gross incomes, their deductions (and exclusions and exemptions), and their tax rates. Only the last of these numbers is written into law (and highly variable law at that), whereas the others depend

\(^{37}\) Note, though, that this is still only a federal calculation. If one were to adopt GA as the preferred accounting framework, it would surely be desirable to extend it to the entire government sector.

on long-term estimates of birth rates, death rates, net immigration rates, productivity
growth rates, homeownership rates, trends in medical insurance coverage by the private
sector, and on and on.

Certainly, some long-term estimates can be quite reliable. Birth and death rates
change rather slowly, so projecting the number of native-born Americans likely to be
living in thirty or forty years is not much of a stretch. On the other hand, other estimates
are notoriously volatile. The CBO has changed its estimates of annual deficits, for
example, by as much as 100% over the space of several months. Even longer-term
forecasts, which are plausibly less prone to temporary blips, are more prone to
cumulative error.\(^39\) Indeed, even history is unstable, as the “New Economy’s”
performance in the 1990’s has been substantially written down.

By far the most important forecast that goes into the generational accounts’
calculations is the time path of labor productivity growth.\(^40\) The Trustees of the Social
Security Administration produces four estimates of this path: the pessimistic scenario, the
mid-range scenarios (two paths that differ only slightly) and the so-called “rosy
scenario.” The Trustees (and GA) tend to default to the mid-range scenarios, choosing
the “moderate” path between pessimism and optimism. On that basis, the long-term
forecasts for the Social Security Trust Fund turn negative in the aggregate in less than
forty years, and the Social Security system runs an annual deficit in less than twenty

\(^39\) The pioneering economic forecaster Otto Eckstein, founder of Data Resources, Inc. (now
DRI/McGraw-Hill) once said to his graduate students: “You can believe our quarterly forecasts rounded to
a full percentage point (e.g., 4.3% growth forecasts mean that growth will be somewhere in the
neighborhood of 4% in the next quarter). You can believe the sign of our annual forecasts. And you
should just ignore our five-year projections.”

\(^40\) Other variables are obviously important, but less so. Choosing an interest rate for the net present
value calculation is essential, for example, and has a highly significant impact on the overall calculations.
See Haveman, \textit{supra} note 33, at 103-04.
years. Similarly, GA relies on these estimates to show that the entire government (including Medicare, which is important because it is a large source of the long-term deficits in these calculations) faces an increasingly unbalanced set of books as we move into the future.

The picture is far less grim than these calculations suggest, however. The “rosy” scenario is, in fact, hardly pie-in-the-sky. Indeed, its projections of long-term changes in productivity growth are rather moderate compared to the last half-century of U.S. performance. Indeed, the rates of growth projected in the mid-range scenarios are below the average of most decades in the twentieth century other than the Great Depression years of the 1930’s. Yet, even with those pessimistic assumptions, only a tiny improvement in productivity growth would wipe out the long-range deficits in the Trustees’ accounts for Social Security.

A recent calculation of this Fiscal Gap, using the most current version of the Generational Accounting methodology, has been provided by Gokhale and Smetters. Admiringly transparent in its description of how the calculations were derived, this study estimates a Fiscal Gap of $44.2 trillion, of which $7 trillion is attributable to Social Security, $36.6 trillion is attributable to Medicare (split roughly equally between Part A and Part B), and only $0.5 trillion is attributable to the rest of the federal government. Gokhale and Smetters provide a range of scenarios under which the Fiscal Gap could be

---

41 Cite and numbers.
43 While the discussion of the mechanics of their estimates is clear, it is hardly dispassionate. Words like “drastic” permeate the discussion, which (given the study’s conclusions) perhaps understandably presents an urgent (even alarmist) tone.
44 *Id.* at 3. (Gokhale and Smetters prefer the term Fiscal Imbalance, but I will use the more common Fiscal Gap.) This is a midpoint estimate. The lower bound is $29 trillion, while the higher bound—“under still quite conservative assumptions”—is $64 trillion. *Id.* at 6.
erased, suggesting as their most likely choice an immediate and permanent 16.6% increase in wage taxes.\textsuperscript{45} Updated estimates based on the same model (crucially including the just-passed Medicare drug benefit) show the Fiscal Gap to have reached approximately $70 trillion. In the discussion below, however, I will focus on the Gokhale and Smetters estimates from 2003, in part because of their clarity, and in part because the analysis will show that the exact numbers provided by such studies are ultimately not helpful in fiscal analysis.\textsuperscript{46}

2. Lifetime Net Tax Rates

Although it is analytically separable from the long-term budgeting calculation, the “generational” part of Generational Accounting is where its political impact becomes most potent. Kotlikoff suggests that it is possible to move from the “What if we did nothing?” question to compare the treatment of different generations based on their lifetime receipts of government benefits and their lifetime tax payments. Using the same method described above, it is possible to choose arbitrary cutoff dates for different generations and then to calculate their “lifetime net tax rates,” i.e., the net present value of their lifetime tax payments minus the net present value of their lifetime government-paid benefits.

Seeming to confirm the suspicion that a large population cohort in a democracy could distort the benefit system in its favor (especially in a democracy in which the young are less likely to vote), Kotlikoff made national headlines when he announced in

\textsuperscript{45} Id. at 6.
1993 that the lifetime net tax rate of younger generations would be 71%, whereas the rate for Baby Boomers would be 35%, and the rate for current retirees was 21%. A few years later, the numbers became even more dramatic, when the 71% figure was increased to 84%.  

B. Weaknesses in Generational Accounting

As appealing as the basic foundations of GA might be, however, the theory does not deliver what it promises. Far from being a neutral tool for dispassionate evaluation by policymakers, the generational accounting model makes the fiscal horizon look far worse than it will probably be, and these results have colored the policy debate for the worse. Moreover, the claim that GA calculations can act as a default early-warning system is at best overdrawn. It is simply not possible to define a clean baseline. Finally, the economic assumptions on which GA is based are too fragile to use for meaningful policy analysis.

1. Paying Down the Debt

The source of the huge differences between generations noted above is quite peculiar. The generational accounts assume that, as of the date that an account is calculated, there are two groups of citizens: the generations that are already born and the one that is about to be born. Then, the accounts compute the taxes that the already-born

---

47 Kotlikoff, supra note 4 (“Deficit Delusion”).
48 Cohen, supra note 6, cites this higher figure, as does Shaviro, supra note 8, at 716. On the other hand, when Kotlikoff recently recalculated his generational accounts and found that the 84% rate for future generations had fallen to 35.81% (cited in Shaviro, supra note 12, at 150), the “good” news was not met with fanfare. (The incredible precision of those estimates is a separate issue.)
49 This means, of course, that the political audience for these estimates, the younger non-voters, was in fact not included in the group that is supposed to be paying nearly all of their lifetime income in net taxes.
will pay minus the direct cash benefits that they will receive (which are in part known, because some taxes and expenditures are already history), under the current tax and spending regimes.

The soon-to-be-born are, however, treated differently. For them, lifetime taxes include not just those that they would be forced to pay under current law, but also taxes sufficient to pay down the entire national debt (accumulated before they were born) during their lifetimes. There is no good reason to assume that the entire national debt will or must be paid in that time, but that is the assumption that drove the dramatic 84% result.

On its own, of course, this assumption cannot help but make things look much worse for the new generation. With an entire lifetime of work ahead of them, and with the government unable to borrow money, they must pay for their own benefits as well as those of their parents and grandparents. The older generations, meanwhile, had a good ride, and they are allowed to continue that ride even while their heirs are paying for the difference.

The generational impact of current fiscal policy, however, is better viewed through the more traditional crowding-out lens. Current deficits are likely to decrease future growth in GDP, which makes future generations worse off than they otherwise would be. The inter-generational comparisons are becoming less relevant as the generations that benefited from the expansion of Social Security and Medicare die off.

50 Haveman, supra note __, at 100.
51 Id. at n.5 (“In effect, there are two implicit fiscal regimes in place during the future years when both members of current generations and members of future generations are living.”)
52 Since the government’s bondholders tend to be older (especially indirectly through retirement funds), this also skew the inter-generational comparison as income is redistributed upward.
Moving forward, the real question is reduced to the now-versus-later question that has always been the central focus of budgetary analysis. We cannot know whether any single future generation will be called upon to pay down the debt; but we can say that any decision that raises deficits at one point in time is likely to cumulatively lower future GDP.

2. Benefits Not Counted in a Generation’s Accounts

The calculation of any particular generation’s lifetime net tax rate also excludes many of the indirect benefits provided by governments—indeed, the very benefits for which governments are traditionally thought to exist. The only benefits that go into the GA calculation are those that are paid in cash. The benefits from cleaner air, pleasant parks, medical research and development, lower crime, etc. are not a benefit. Taxes pay for them, but they are a net cost of government in the GA calculations.\(^{53}\)

It is not clear \textit{a priori} how this fault in the generational accounts would affect inter-generational comparisons. Indeed, it is imaginable that these benefits are so diffuse that they benefit every citizen equally. It is also possible, though, that some of these benefits are disproportionately shared. The cost of educating the Baby Boomers was borne by our parents, yet all future generations will benefit from it.

Leaving that very open question aside, though, the fact that the GA calculations of lifetime net tax rates are skewed upward is important simply because it skews the political response. If the members of Gen X are told that their net tax rate is 84\%, while that of their parents is 35\%, they are likely to have two responses: 1) Our generation is

\(^{53}\) It is, of course, always possible to adjust the GA calculations to take these non-cash benefits into account. To the extent that this can be done, GA begins to resemble capital budgeting—which is all to the better. See below.
being cheated, and 2) *All* generations are being cheated! After all, while 35% is better than 84%, paying more than a third of your lifetime earnings to a government that (according to this model) does not do anything useful with the money is pretty upsetting.

If, on the other hand, the numbers were 8.4% and 3.5%, the magnitude of the inter-generational backlash would be muted (since outrage is likely to be at least partly based on the magnitude of the difference as well as the proportional comparison), and the anti-government reaction might not even register politically.

In other words, the effects on society of this widely-quoted statistic go beyond the simple, modest claim that GA is just a diagnostic tool. It is a political tool, and its affects are predictable.

### 3. Demographic Trends

The most important demographic phenomenon facing the U.S. and other Western countries is the Baby Boom and the subsequent dramatic decline in birth rates after 1964. While this phenomenon will end within a few decades, at least some parts of the long-term GA deficit are not going to be solved simply by the death of the Boomers.

The more important trend is the general increase in life expectancies over the long term. Indeed, given long-term trends in health, even after the 75-year window, the paths of receipts and expenditures continue to diverge, as an increasingly large non-working aged population consumes more of the economy’s resources, largely through the health care system.

If this is accurate, it would mean that the focus should not be on the effects of the Baby Boomers but on reining in our seemingly insatiable appetites for medical care.
While that might be a wise policy on its own, however, the GA framework does not provide a compelling reason to adopt such limits.

For example, Shaviro cites research indicating that healthcare expenditures on the elderly will continue to rise significantly relative to GDP. This, however, assumes that health care spending on an aging population will show the same trends as current health care expenditures. For example, if the typical 75-year-old today consumes $X of health care, and if there will be twice as many living 75-year-olds in 50 years, then it would appear that health care spending would have to double in the future. This assumes, in turn, that even though life expectancies will rise, elder health at specific ages will not improve.

We know, however, that a large fraction of the money spent on health care is spent at the very end of life—heroic, life-prolonging procedures that add a few weeks or months to the lives of chronically ill patients. If those chronically ill patients do not become chronically ill until twenty years later in life, however, there is no reason why overall health care spending must rise—even if we fail to change the way we deal with end-of-life decisions.

This is not to say that it is impossible to imagine a future with higher health care expenditures. It does indicate, however, just how difficult it is to rely on estimates of health-related spending decades in the future.

The calculations in Gokhale and Smetters, on the other hand, are based on a much simpler assumption, that medical care will grow for the next 75 years at a rate one percentage point faster than the growth rate of GDP, then fall over the ensuing 20 years

54 Shaviro, supra note __, at 152.
to grow at a rate equal to GDP growth. These assumptions mirror those made by the Medicare Trustees.\textsuperscript{55} The significance of these assumptions is considered below.

C. The Policy Imperatives

The policy regime that GA calls for is clear: Fiscal contraction, austerity, and pain. The positive spin is that the pain can be “shared.” Two somewhat inconsistent norms are offered in designing policies to redress the long-term imbalances.

1. The Norm of Generational Equality

The most obvious norm, which Kotlikoff explicitly relies upon, is simple generational equality.\textsuperscript{56} Every generation should pay a lifetime net tax rate no higher than the last. Other than symmetry, there is no apparent philosophical imperative behind this norm. Indeed, since the whole notion of lifetime net tax rates tells us nothing about real living standards (net of those tax rates), there is no obvious reason to be concerned about lifetime net tax rates \textit{per se}.

2. The Norm of Shared Sacrifice

Recognizing this, Shaviro argues for the norm not of equal tax rates but of shared sacrifice. “With respect to generational distribution, once the members of a given age cohort have died, they can no longer be asked to share in the pain of tax increases or benefit reductions.”\textsuperscript{57} In other words, it is important to take the benefits away from the old people now because they are alive now. While there are certainly colorable (but

\begin{footnotesize}
\textsuperscript{55} Gokhale and Smetters, supra note \_, at 23.
\textsuperscript{56} See Diamond, \textit{supra} note \_, at 1.
\textsuperscript{57} Shaviro, \textit{supra} note \_, at 155.
\end{footnotesize}
highly debatable) arguments that the elderly over-consume,\textsuperscript{58} the current elderly also have a strong argument that they earned it.

\textbf{D. Generational Accounting Without Generations?}

Populist appeal aside, the importance of the GA approach should not be underestimated. Kotlikoff and his co-authors argue that they have created a meaningful baseline, allowing us to assess the effects of any proposed government policy over the long run, and potentially comparing that affect on different age groups. If these claims were true, then it would be important for everyone to send their policy proposals through the GA machine. If the results of such an analysis turned out negative, the policy would be presumed harmful until proven beneficial.

On the other hand, it is also possible to take the generations out of generational accounting, by simply stopping after we compute the Fiscal Gap. We can look at long-term trends and calculate long-term tax revenues and government spending assuming a continuation of current policies. If this calculation shows a deficit, then it is a warning: Unless something is changed, the government will have to cut spending or raise taxes on future generations. While we might still choose to do nothing, the argument continues, at least GA gives us a fair number to work from.\textsuperscript{59}

\textbf{1. Another Baseline Problem}

The supposed agnosticism behind GA calculations thus rests on the idea that they are warnings, not predictions. Indeed, if GA is used as its proponents suggest, we enable

\textsuperscript{58} See Cohen, \textit{supra} note __.
\textsuperscript{59} Even Haveman, \textit{supra} note __, despite offering a withering criticism of GA, allows that this “base case” analysis can provide useful information. \textit{Id.} at 100, 110.
ourselves to make “tough but sensible” choices now to prevent the disaster that surely awaits us if we fail to act.\textsuperscript{60} In fact, though, the theory is not agnostic, and its proponents’ call to action rests on weak assumptions.

The underlying question is one of default. Once the cross-generational comparisons are set aside, GA claims to ask, in essence: “If we were to change absolutely nothing, and if we can believe the forecasts on which our estimates are based, are our current fiscal policies allowing living generations to steal economic resources from their grandchildren and their grandchildren’s grandchildren?” But it does not stop there. If the answer to that question is yes, then the argument is that we must act now. Why now? Because our current path is unsustainable, and delay is only going to allow matters to get worse.\textsuperscript{61} Sacrifice today means sowing a greater harvest tomorrow.

While this logic surely appeals to our Puritan roots, it is misleading. We should only enact new policies now if we believe that we will not change policies along the way for other reasons. By way of analogy, consider the argument (known as “bracket creep”) that tax rates in the 1970’s would have led to an ever-higher percentage of GDP being collected by the IRS, because inflation was pushing everyone inexorably into the highest tax brackets. If that argument had been coupled with a call to raise government spending immediately, because we can count on ever-higher revenues in the future, then surely that would have been foolish. Everyone knew that Congress would pass regular tax cuts to undo the effects of bracket creep; so even if no one anticipated the indexing of brackets to

\textsuperscript{60} Shaviro, \textit{supra} note __, at 716 (arguing that politicians should “openly face[s] today painful choices that ultimately will have to be faced anyway.”)

\textsuperscript{61} \textit{Id.} (arguing that we should act “sooner rather than later”). See also Gokhale and Smetters at 3.
inflation in the 1981 tax bill, certainly no one thought that everyone would end up in the 70% tax bracket when inflation had made even a street-sweeper “rich.”

Similarly, the GA call to action loses much of its appeal when we realize that this “interesting calculation” is based on some extremely arbitrary economic forecasts combined with the arbitrary assumption that only a few things are set in stone. Consider the current path of tax rates. Current tax law is in a bizarre state, because of the ten-year reversion feature of the 2001 tax bill. The estate tax is set to decline, disappear, and reappear. The 28% tax bracket declines to 25% and then returns to 28% (with similar moves in the other brackets, and some brackets merging into others only to reappear in 2011). No one expects the reversion to happen. The House in 2002 passed, on partisan lines, a bill to make the 2001 cuts permanent—although no one viewed that action as anything more than election year posturing. Even with the current one-party dominance of the federal government increasing the likelihood of tax cuts, it is not clear what form those cuts will take.

With the status quo ante not a serious possibility, and the status quo post unlikely ever to become the status quo at all, how can one even formulate a call to action on the basis of a GA calculation? In addition, current projections assume that the Alternative Minimum Tax (which is not indexed to inflation) will remain in place, even though it will

62 There are, of course, strong political economy arguments for indexing the tax code rather than relying on ad hoc corrections; but those are beside the point here.
63 See Haveman, supra note __, at 99 n.4.
64 Because of a self-imposed super-majority voting requirement in the Senate, bills that would reduce tax revenues beyond ten years from the effective date of a bill require sixty votes to pass. Lacking sufficient votes, the Senate made all provisions of the 2001 tax act void in 2011.
almost surely be altered or repealed when it begins to affect large sections of the middle class.\textsuperscript{65}

Carrying this over to the demographic argument is potentially even more devastating to the GA position. If the time paths of revenues and expenditures do not come back together after the death of the Baby Boom generation because of the trends in life expectancies, then surely it is important to forecast how those longer life expectancies will change both individual behavior and government policy. As people live longer, they will naturally use more economic resources, including medical care. (They might also produce more economic resources.) Based on current law in which retirement ages are rising to 67 and then staying put, GA estimates would have us believe that in 75 years—assuming that all of the other forecasts over that time span are true—we will have a nation of impressively healthy septuagenarians (and older) living off of the sweat of a relatively tiny population of younger workers.

Why make the assumption that the retirement age (effective if not statutory) will remain fixed as the population inevitably ages healthfully? Boredom alone is likely to lead to a changed politics of retirement. While an advocate of GA can always claim that their calculations can tell us what happens if retirement ages do not change, that is a far cry from justifying the argument that we need to cut benefits and raise taxes \textit{today} on Medicare beneficiaries in order to bring the long-term budget into balance.

Indeed, Gokhale and Smetters do allow themselves to make a single departure from their blanket assumption that policy is set in stone. Separate from the “\textit{bracket creep}” caused by inflation, real economic growth can cause “\textit{real bracket creep},” whereby

increases in real income cause every taxpayer’s income ultimately to rise to the highest bracket. Gokhale and Smetters quite reasonably view this is absurd and thus assume that the brackets will be adjusted over time to prevent this from happening. This, of course, makes the estimated Fiscal Gap look worse; but more importantly, it raises the question of why this is the only concession to reality that is allowed in the estimates going forward.

2. The Fiscal Gap and the Annual Deficit

The absolute size of the estimated Fiscal Gap is important for another reason. “[T]he [Fiscal Gap] grows by about $1.6 trillion per year to $54 trillion by just 2008 unless corrective policies are implemented before then. This rapid annual increment is also about ten times as large as the official annual deficit reported for fiscal year 2002.”\textsuperscript{66} This suggests that the “true” deficit is not measured accurately by the annual cash-flow deficit but by the change in the Fiscal Gap from year to year. Because that annual change, in the absence of policy enactments, is simply equal to the previous year’s Fiscal Gap times the assumed annual interest rate, this annual quasi-deficit measure will be much smaller if we allow the forecasts in the Fiscal Gap to include changes in retirement ages, health care cost trends, etc. If the estimated Fiscal Gap were only $15 trillion, in other words, the annual quasi-deficit would be just over $500 billion, whereas a Fiscal Gap of $88 trillion would imply an annual quasi-deficit of $3.2 trillion. The annual changes depend completely on the accuracy of the aggregate estimate.

\textsuperscript{66} Gokhale and Smetters, at 3.
3. Making the Fiscal Gap Disappear

Even if it were possible to agree on the likely future path of policy decisions, the Fiscal Gap can be manipulated simply by enacting policies which will take effect in the future. If, instead of Gokhale and Smetters’s suggested immediate increase in wage taxes of 16.6%, Congress enacted an increase in wage taxes that started at 0.1% in twenty years and rose to some level well in excess of 16.6% twenty years later, the Fiscal Gap would immediately become zero. Cynics would argue that this is non-credible, and they would have a point. As enacted policy, however, such a law would make the Fiscal Gap equal to zero, by definition. The annual quasi-deficit would then also equal zero. While nothing would have changed, these fiscal measures would show nothing amiss. It is true that there would still be a relative burdening of future generations, but not one that has changed because of the adoption of the new law.

4. How Far Into the Future?

Finally, Gokhale and Smetters use an infinite horizon rather than the 75-year horizon that is the norm in such analyses. As they point out, their model predicts a Social Security gap of $1.6 trillion, compared to an infinite-horizon gap of $7 trillion, while the 75-year Medicare gap is $15.1 trillion, compared to the $36.6 trillion in their infinite-horizon model.\(^67\) In other words, over sixty percent of the Fiscal Gap occurs from 2078 through infinity.

Given that the bulk of the estimated Fiscal Gap, both pre- and post-2078, is caused by the Medicare growth assumption, it is probably more accurate to describe any

\(^67\) Gokhale and Smetters, at 34.
long-term fiscal crisis as a health care crisis. If the cost of medical care continues to grow in future decades at rates exceeding the growth rate of GDP, then certainly a policy intervention will become necessary. Prescribing policy initiatives for current lawmakers on the basis of such broad assumptions (even if those assumptions are arguably “conservative” by recent historical standards) borders on being arbitrary.

In short, the Generational Accounts are not a useful guide for policy, because they are based on highly unreliable forecasts (and those forecasts are unnecessarily pessimistic), they ignore the likely path of political decisions, and they are far too easy to manipulate. Moreover, they cannot meaningfully compare the relationships among generations in how they share the cost of running the federal government, at least in a way that is different from the standard crowding-out approach. While the exercise of measuring Fiscal Gaps is based on a reasonable desire to see what we might be getting ourselves into, the mechanisms for such long-term forecasts are simply too crude to add meaningfully to our arsenal of policy choices.

---

68. Paul R. Krugman, Social Security Scares, The New York Times, Mar. 5, 2004, at A23 (“The projected rise in Medicare expenses is mainly driven not by demography, but by the rising cost of medical care, which in turn mainly reflects medical progress, which allows doctors to treat a wider range of conditions.”)
69. Arguably, the lesson to be drawn from an exercise like that of Gokhale and Smetters is that the news is good, i.e., outside of the more general question of how to handle health care costs (both inside and outside of Medicare), we do not apparently face significant long-term budget problems. If so, then the exercise is worthwhile in a negative sense. Still, there are less favorable assumptions that could make the Fiscal Gap calculation look worse, and those assumptions are subject to all of the uncertainties described here. If Fiscal Gap calculations are to be but one piece of information among, therefore, they at least should not be given the prominence that Gokhale and Smetters would give them.
III. Capital Budgeting

If Generational Accounts were the only available method of making decisions about long-term fiscal policy, then it might be valuable to try to nail down some reasonable (though still arbitrary) long-term budget projections. In fact, though, it is not necessary to look at generational accounts at all in assessing the impacts of spending decisions.

A. Capital Budgeting

For legal and policy analysts, the most important issue in assessing any government policy is not the effect of the policy on the deficit (properly measured) but rather the question of whether the policy is “worth it.” This can only be understood from the perspective of capital budgeting.

1. Definitions

A capital budgeting system separates expenditures into two categories, operational and capital. The operating budget accounts for the purchase of goods and services whose full benefits are enjoyed during the year in which they are made. The capital budget accounts for items whose benefits are longer-lasting. One rough estimate of the fraction of federal spending that can be categorized as capital expenditures is 25%. 

Many capital projects are likely to bring with them the requirement of at least a minimal level of maintenance expenditures. Depreciation on the existing capital stock and maintenance expenditures are thus netted out of the capital account.  

See, for example, Neil H. Buchanan, Debt, Deficits, and Fiscal Policy Three Essays, Ph.D. dissertation, Harvard University, 1996, and citations therein.
Far from being an innovation, capital budgets are the norm virtually everywhere but in the federal government. Publicly held corporations must separate operating and capital expenditures, by accounting convention. Indeed, most “profitable” corporations would not be viewed as profitable if they were prevented from segregating their capital expenditures, since even the most profitable corporations borrow money every year (that is, they run “deficits”).

Similarly, as noted above, state and local governments overwhelmingly use capital budgeting. The oft-noted fact that most U.S. state governments operate under balanced budget requirements is, indeed, not what it seems, because a state can still borrow money for capital spending even if it does not have the tax revenue to pay for it.  

It is odd, therefore, that the federal government would not use capital budgeting. The arguments against adopting capital budgeting are prudential, asserting that it is simply not wise to trust Congress with such a powerful tool for justifying deficit spending. With a capital budget available, the argument goes, any silly expenditure can be slipped into the federal budget and camouflaged as capital spending. The entire process, this argument continues, is open to abuse and gamesmanship.  

Indeed, former President Clinton drew fire from the nation’s English teachers when he attempted to describe some spending policies as investments. “[Clinton] captured third place in the 1993 Doublespeak Award, administered by the National Council of Teachers of English,

Note that, even with a capital budget, the perversity of failing to cyclically adjust remains. If a state is required to equate actual tax receipts with actual expenditures, balance will become deficit when the economy goes South—and the downward spiral will continue as long as the state’s politicians are required to balance their non-cyclically adjusted operating budget.  

See, for example, Karen Pennar, Beware of Accounting Magic Tricks, Mr. Clinton, Business Week, January 18, 1993, 55.
for . . . his insistence on ‘using the word “investment” as a substitute for the word “spending” in his rhetoric on economic policy.’”74

While it is certainly possible to over-use the word “investment,” the English teachers simply had their accounting wrong. The choice is not between spending and investment, but between investment spending and consumption spending. If we view government as having to choose between the two, then our goal should not be to prevent politicians from using the term “investment,” but rather to ensure that they use it correctly.

2. Using Capital Accounts

While it hardly stretches the imagination to suspect that members of Congress do not always act in accordance with pure economic theory, it is simply not true that a capital budgeting process is an open cookie jar. Accounting standards have been promulgated and are taught in every business school in the country. States and corporations do, of course, sometimes play at the edges of these rules,75 but the rules can and do constrain behavior.

Moreover, the current federal system effectively treats all expenditures as if they were operating expenditures. While it is possible to adopt a capital budget and then to implement it incorrectly, not to implement it literally guarantees that the government’s budget is measured incorrectly. It also leads to poor policy choices. If the only goal is to balance the budget, cutting projects with valuable long-term payoffs looks just as good as


75 The reader is encouraged to provide her own joke regarding Enron, WorldCom, Tyco, etc. here.
cutting pure pork. And it is even “desirable” to sell public assets at a loss, since any
revenue received reduces the annual deficit.

Separating investment from consumption, therefore, has at least two desirable
effects. First, it would prevent policymakers from cutting programs that are likely to
provide long-term benefits to the economy. For example, to prove their fiscal
responsibility, the leadership of the House of Representatives is intent on cutting funding
for mass transit, because they “see transportation projects as one of the first ways to cut
back the budget and reduce the deficit.”

While it is plausible that some transportation projects would not provide long-
term payoffs, others surely would. In future research, I will explore the administrative
machinery that might be put in place to allow policymakers to engage in reasonably fact-
based inquiries into the likely payoffs of various spending projects—without providing
room for budget mischief. For present purposes, the assertion is simply that there are
some projects that would qualify as capital spending and that it would be wise to fund
them even from borrowed funds. At the very least, it remains to be proved that the
political gamesmanship that would exist under a system of capital budgeting would
necessarily be worse for the economy than our current system.

The second advantage of adopting a capital budget for the federal government is
that it would penetrate the category commonly known as “pork-barrel” spending. The
operative definition currently seems to be that pork is anything that directly benefits a
specific legislator’s constituents. For example, Senator Robert Byrd (D., W. Va.),

———

76 Raymond Hernandez, Senate Panel Backs Transit Aid for New York, The New York Times,
probably describes himself as the “prince of pork” for his ability to direct federal projects into his home state. Among his successes:

- There are two Robert C. Byrd United States Courthouses,
- four Robert C. Byrd stretches of highway, freeway, expressway and drive, and a Robert C. Byrd Bridge. And two Robert C. Byrd Interchanges to reach these valuable amenities. There is the Robert C. Byrd Lifelong Learning Center, the Robert C. Byrd Hardwood Technology Center, the Robert C. Byrd Health and Wellness Center, and the Robert C. Byrd Institute for Advanced Flexible Manufacturing.\(^77\)

Again, while it is possible that each of those items is actually wasteful, at least their titles bear the earmarks of projects which state governments classify as capital spending projects. Going beyond that superficial level to identify the valuable projects would be an important benefit of adopting a capital budget.

### B. Beyond Capital Budgeting

Clearly, this goes beyond simply arguing in favor of adopting a system of capital budgeting. It also suggests that even balancing an operating budget is not sufficient. Balancing the operating account, sometimes called the Golden Rule of Budgeting,\(^78\) actually allows the government to waste resources if it is able to raise the taxes to do so. This should be unacceptable. Instead, the government’s decisions should always be driven by considerations of whether those decisions are helpful to the current and future health of the economy.


Dealing with Medicare is a good example. As Shaviro notes, many Medicare expenditures are likely to improve the health of seniors. This might then plausibly reduce the net cost of health care even as life spans lengthen. If so, then we might not wish to cut Medicare benefits today on the basis of our concern for future generations. At the very least, as noted above, a more careful analysis of the nature of long-term health care trends is necessary.

C. Back to Small Decisions

An unexpected benefit of piecemeal capital budgeting is that we would not be forced to rely on aggregate macroeconomic forecasts of the sort required by Generational Accounting. Case-by-case analysis actually makes more sense than aggregate analysis in this framework, because it is possible to say that a specific policy (such as funding early-childhood nutrition programs) is likely to achieve certain goals over time (or not), based on a much more manageable economic forecast. Being wrong about one policy need not cause us to be wrong when estimating the effects of other policies.

IV. What Does it Mean to Be Fair to Future Generations?

The unspoken assumption in all of this discussion, not just of capital accounting but of generational accounting as well, is that the government should never make decisions that would reduce economic growth in the future. Even if one completely agrees that the measurement problems discussed above must be addressed, and that there should be a capital budget, one is still proceeding from the assumption that the bad decision is the one that would reduce the total capital stock that current generations bequeath to the future.
A. How Much is Enough?

It is possible, as noted earlier, that even a decision by the government to consume economic resources rather than to invest them will not harm future economic prospects. This will happen if the resources that the government purchases for its use would have been used to produce private consumption rather than private investment. Conceptually, if the government throws a wild party by hiring the people who would have worked at a privately-funded wild party, then future generations are unaffected. It is only if the government’s party is (indirectly) staffed by computer programmers and construction workers that the future productive capacity of the economy is reduced.

This, however, still leaves open the question of why we must maximize the capital stock that we pass on to our heirs. Given that economic growth is generally on an upward trend, why is it necessary to give our wealthy grandchildren even greater wealth? The bipartisan (within the economics profession as well as among politicians) silence on this question is notable, to say the least. While there have been preliminary attempts to estimate how much capital should be produced for future generations, the unspoken assumption is quite blunt: We cannot do anything to reduce the capital stock that we bequeath to our children and grandchildren. Perhaps it is time to question that assumption more aggressively. (Moreover, although the subject for a different essay, it is equally important to account for the “intergenerational unfairness” created by problems such as environmental damage.)

---

B. The Real Inter-Generational Issues

The very language of intergenerational transfer is, therefore, potentially misleading. Indeed, it is not possible to “pass the bill to future generations” for our current spending. When the government uses economic resources, the rest of the economy currently cannot use those resources. (This, of course, assumes that those resources were going to be used at all. Given the prolonged slack in the global economy, even that assumption is often contestable.) That means that we pay for what we do, in the fundamental sense of opportunity cost.

Future generations are, of course, affected by these decisions, too. If, as discussed earlier, the government’s decisions are likely to decrease the net capital stock that is passed on to future generations, then their output will be lower than it would otherwise be. Therefore, the best approach is to think about how the government is using current resources. If it is investing them, then future generations will benefit. If it is consuming them (or simply wasting them), then they will not.

Even more fundamentally, it is not at all obvious that cutting benefits to seniors today will hurt only seniors. When the elderly lose benefits, they can turn to their children to make up the difference. Even if they do not do so directly, they can consume more of their estate than they otherwise would have, thus reducing the wealth of their children. This incidence question indicates just how difficult it is to measure meaningfully the impact on different generations of our fiscal policies.
V. Conclusion

The traditional debate about budget deficits witnessed a divergence between the economic analysis, which saw that deficits are poorly measured in the U.S. and argued that certain deficits are beneficial for the economy, and the political view that every deficit is evidence of moral failure. This unusual stalemate is currently on hold, as the brief era of surpluses gave way to the (hopefully even more brief) era of terror, leading to a decreased emphasis on fiscal orthodoxy.

In addition, an alternative approach to budgeting, Generational Accounting, has emerged. Designed to correct some of the weaknesses of annual budgeting, GA purports to provide an “early-warning system” to allow us to correct our long-term fiscal imbalances before it is too late. Unfortunately, this theory is based on highly contestable assumptions, makes questionable analytical choices, and is inherently incapable of providing the useful baseline that its proponents promise.

Instead, a modified system of capital accounting should be used to guide economic policy. This would emphasize case-by-case analysis, allowing legal analysts to compare the likely costs and benefits of policy proposals while keeping a clear eye on the importance of government investment in our future prosperity. If political concerns about the potential abuse of capital budgeting prevent the federal government from adopting an explicit capital budget, the best response would be to continue to rely on the current (admittedly imperfect) budget measures, which at least provide some useful guidance regarding the immediate effects of our fiscal policies.
Exploring the Treatment of Borrowing and Accruals in a Realization Income Tax

Joseph M. Dodge*

[Draft of March 2, 2006]

Income tax theory refers to those aspects of the income tax base that derive from the concept of a taxpayer’s “net increase in wealth” over the taxable period (the taxable year). This article examines the concept of “decrease in wealth” as it pertains to borrowing and liability accruals under two differing income tax models, the ability-to-pay (ATP) model and the Simons income tax1 model,2 and their mongrel offspring, the realization income tax.

A basic feature of the U.S. income tax from its 1913 inception is the so-called “borrowing exclusion,”3 under which borrowed cash is not included in the gross income of the borrower on the ground that the borrowed cash, an increase in wealth, is instantly offset by a corresponding decrease in wealth represented by the repayment-of-principal obligation. Another longstanding feature of the U.S. income tax has been the rule whereby taxpayers using the accrual method of tax accounting have been allowed to claim current deductions in an amount equal to the face amount of future obligations to make (deductible) payments (typically as goods and services are acquired on credit).4 Both of these features are in essence the same: the tax base is currently reduced by the face (or stated principal) amount of obligations to make future payments.

The thesis of this article is that these features, unwittingly borrowed from business accounting, are worthy of scrutiny. Specifically, I explore the possibility that debt and liabilities might, depending on one’s vantage point, not be viewed as involving a current decrease in wealth under the income tax. This exploration is carried out by posing the possibility of an alternative set of tax treatments for borrowing and liabilities, which I call the “principal inclusion/deduction” (“PID”) approach. Liabilities arise in

* Stearns Weaver Miller Alhadeff & Sitterson Professor, Florida State University College of Law. The author wishes to thank Calvin H. Johnson for helpful suggestions.

1 The Simons income tax is developed in Henry C. Simons, The Personal Income Tax (1938). Simons acknowledged that the core concept of income “net increases in wealth plus consumption over the taxable year,” see id. at 50, was anticipated by the earlier commentators Georg Schanz and Robert Haig. Hence, the core concept of income is often referred to as “Haig-Simons” income or “Schanz-Haig-Simons” income. However, Schanz work (to my knowledge) has not been translated from the original German, and Haig did not elaborate much on the concept, so that Simons is by far the leading authority in this respect.

2 The two concepts of income also differ with reference to off-the-bottom (subsistence-type) allowances, which are integral to the ATP concept but exogenous to the SIT concept (although Simons, supra note 1, at 220, acknowledged the necessity of them).

3 This treatment appears to be common to all income tax systems worldwide. See Hugh J. Ault & Brian J. Arnold, Comparative Income Taxation: A Structural Analysis (2d ed. 2004) (containing no discussion of the borrowing exclusion).

4 Taxpayers can instead use the cash method of accounting, unless the accrual method is required by the Internal Revenue Code and Regulations. See I.R.C. § 448 (limiting use of cash method); Treas. Reg. § 1.446- ?? (inventories are to be accounted for using the accrual method). Most business activities follow the accrual method, and individuals mostly use the cash method for non-business activities. See I.R.C. § 446(a) (tax accounting method usually follows method taxpayer follows in keeping books).
various ways, so that it is hard to describe the operation of the PID approach in every circumstance in one sentence. Nevertheless, it can be said as a crude generalization that the PID approach is essentially a cash-method approach, in contrast to the prevailing accrual-method approach. Thus, in the case of cash borrowing, the borrowed cash would be included in gross income, and deduction should be obtained for loan principal repayments and, in the case of business and investment borrowing, interest. In the case of liabilities that are not the equivalents of cash borrowings (that is, liabilities to pay expenses), any deductions would be taken only when the liabilities are paid. The common-sense justification for the PID approach is that a decrease in wealth in the tax sense should not be deemed to occur until payment is actually made, because prior to payment the taxpayer has the use and control over the funds.

Part I critiques the present system and describes the operation of the PID approach.

Part II examines the doctrinal and administrative pros and cons of the PID approach.

Part III examines the issue of whether the PID approach would satisfy the norm of minimal economic efficiency, mostly with respect to debt-financed investments. The heuristic here is an investment that breaks even before tax, which should neither generate a tax liability nor generate a loss that would shelter other income from tax. Part III also examines how debt-financed consumption would be taxed under the PID approach.

Part IV focuses more narrowly on non-borrowing-related liabilities, that is, liabilities to pay future expenses. This requires examination of the merits of the cash and accrual methods of accounting in terms of the “realization” principle. It is concluded that the cash method is to be preferred in virtually all cases.

Part V attempts to pull the various strands together and posit some conclusions that might be drawn from the analysis. In the end, there is no blanket endorsement of the PID approach, but nevertheless a partial PID system may be worth considering.

I. HOW FUTURE PAYMENT OBLIGATIONS SHOULD BE TREATED UNDER INCOME TAX THEORY

This Part has a doctrinal cast. First, the conventional wisdom is reviewed. Second, the conventional wisdom is critiqued from the ability-to-pay (ATP) point of view. Third, arguments are advanced for an alternative approach to the tax treatment of future payment obligations (referred to as the principal inclusion/deduction (PID) approach.

A. The Conventional Wisdom

Under the Simons concept of income (the SIT), which has been the gold standard in academic circles since the 1940s, “income” is stated to be a taxpayer’s net increase in
wealth plus consumption over the taxable year. The proposition that changes in the net wealth of a taxpayer should be included in the tax base is ultimately founded on the ATP norm, which is principally a norm of substantive tax fairness that supplies an answer to the question of “how should a given tax societal tax burden be apportioned among the population?” Simons himself avoided hanging the SIT on the hook of the ATP norm as he knew it, which was based on utility theory. Nevertheless, the SIT as developed by Simons and his followers can be said to embody the notion of “objective” ATP as measured by accounting, through market outcomes, for the material resources commanded by individual taxpayers, with the obvious implication that such resources as are taken through the income tax are available for redistribution by way of government programs. Stated simply, wealth, whether retained or dissipated through consumption, represents ability to pay during the taxable period.

Nevertheless, it would be error to simply equate an ATP income tax with the SIT. The SIT can be described as that modification of a “pure” ATP tax as will (disregarding other government activities) achieve (if not maximize) economic neutrality. Central to any version of the income tax is the maxim that the “same dollars” can only be taxed once to, or deducted by, the same taxpayer. Implementation of this maxim avoids the possibility that the tax can be greater than the net profit or gain of a taxpayer over time. Stated more formally, and assuming a tax rate that is less than 100% of net income, an income tax assures that a net before-tax profit will produce a profit after tax. Moreover, the “ideal” version of the SIT income concept keyed to all (realized and unrealized) changes in wealth (expressed in terms of market values) results in neutrality across

---

5 The classic statement is found in Henry Simons, Personal Income Taxation 50, 206 (1938): “Personal Income may be defined as the algebraic sum of (1) the market value of rights exercised in consumption and (2) the change in the value of the store of property rights between the beginning and the end of the period in question.”

6 It is internal to tax because it does not itself deal with the issue of rates (apart from subsistence allowances “off the bottom”) or the manner in which tax revenues should be spent. See Joseph M. Dodge, ???.

7 See Simons, supra note 1, at 5-17.

8 Simons, supra note 1, largely avoids discussing the normative foundations of his concept of income. Nevertheless, the gist of his system (although not without considerable equivocation) seems to be a reliance on market outcomes, as opposed to subjective utility, see id. at 50-55, with market outcomes being a baseline for redistribution, see id. at 31 and 41. Simons was also concerned that the tax system “not discriminate” among taxpayers. See id. at 106-109. This concern for horizontal equity, rather than economic neutrality, seems to have driven Simons to favor taxation of imputed income from consumer durables. See id. at 112.

9 Henry Simons himself did not deploy the ATP norm as support for the income concept, because he wanted to avoid defective versions of the ATP concept, namely, the “faculty” notion and the “subjective ATP” (utility-based) version.

10 A “pure” ATP tax could be conceived of an annual wealth tax, entailing multiple taxation of the same dollars over time, combined with a separate consumption tax. A dual-tax ATP system would require different rate schedules for each tax.

11 Economic “neutrality” is a necessary condition (along with free markets) of economic efficiency. It is commonly claimed that a cash-flow consumption tax would be better than an income tax with respect to economic efficiency. See Bank, am, Shaviro ???. The counter-claim is that an income tax is the best trade-off between fairness and efficiency. Musgrave ???

12 The single-taxation notion is not unique to the income tax, however, as it is also a characteristic of a cash-flow consumption tax (referred to herein as the “CT”). However, the CT is not based on the ATP norm, because unconsumed wealth is omitted from the tax base.
investments in the sense that after-tax investment returns will bear the appropriate relationship to before-tax returns.  

A major problem in implementing an income tax is that of when and how to take increases and decreases in wealth into account. A thesis of this article is that the ATP norm and the ideal SIT dictate different approaches to the taxation of borrowing and liabilities. The situation is further muddled on account of the fact that the current income tax incorporates the “realization” principle, under which changes in wealth are (generally) taken into account only when “realized.” However, there is no universally-applicable definition or concept of realization. Rather, there is an assortment of rules that can be described as pertaining to realization. Really, any rule or doctrine that pertains to when changes in wealth are taken into account for income tax purposes can be characterized as a “realization” rule, although in practice a realization rule is thought of as one that defers taking a wealth change into account relative to the “correct” moment. Of course, what is the correct moment can itself be controversial. In any event, the inconsistency of realization rules is cause for concern, and this is nowhere more evident than in the case the tax treatment of borrowing and liabilities.

The realization principle is usually said to arise out of the practical consideration that the changes in the fair market value (FMV) of wealth that would be reckoned under an ideal SIT are often hard or impossible to ascertain. It happens that the practical problem of difficulty of valuation converges with the ATP value of liquidity, because hard-to-value assets are often illiquid and vice versa. Under the ATP fairness norm illiquid wealth does not represent “ability to pay,” and should be excluded from the tax base. Although valuation and liquidity problems tend to be overrated, the realization principle has persisted due to its parallelism to accounting, the popular notion that unrealized gains and losses are temporary and hence “unreal,” and inertia. In addition, both difficulty of valuation and illiquidity implicate privacy and autonomy concerns, as in the case of human capital, closely-held business interests, and tangible assets possessing a sentimental value. These issues are resolved (in various ways) by way of “realization” rules. However, realization rules produce disparate treatment of investment returns across categories of assets, and that in turn undermines economic efficiency.

---

13 Thus, in an ideal SIT (unconstrained by practical considerations), depreciation adjustments (whether negative or positive) should reflect changes in the fair market value of assets. See Paul A. Samuelson, Tax Deductibility of Economic Depreciation to Insure Invariant Valuations, [1964] J. Pol. Econ. 604, 606 (1964).

14 The realization principle extends to publicly-traded securities (etc.), which are neither hard to value nor illiquid. In other cases, proxies to valuation can be found. Cite Shakow ???. See generally Scoles and ???.

15 Traditionally, accounting has avoided valuation. The realization idea is a technique of avoiding valuation.

16 Cite Bank ???

17 The most salient realization rule, I.R.C. § 1001(a), simplifies income accounting by reckoning gain or loss with respect to assets only on the occasion of the disposition of assets.

18 Assets that produce returns in the form of periodic cash payments are treated differently than assets that produce returns in the form of appreciation. Certain commentators favor a cash-flow consumption tax precisely on the grounds that the realization principle corrupts the SIT. See Andrews, Shaviro, ??
The subject of borrowing and liabilities comes under the “change in wealth” component of “income” in all of its versions. The conventional wisdom has been that the act of borrowing money cannot give rise to income, because the borrowed cash (which is positive wealth) is fully offset by the liability to repay the loan (which is negative wealth).19 In other words, borrowing cannot be an “income” event because it does not give rise to a current increase in wealth.20 In the same vein, repayments of loan principal are not deductible, because the loss of cash is matched by an equal reduction in negative wealth (the principal liability).21 The approach just described was (and still is) in full accord with business accounting principles.22

This treatment of borrowing has been so ingrained in the income tax that it was thought not to be necessary to enact a statutory borrowing exclusion.23 In basic accord with the borrowing exclusion has been the rule that business taxpayers using the accrual

---

19 See, e.g., Marvin A. Chirelstein, Federal Income Taxation 46-47 (10th ed. 2005) (no increase in taxpayer net worth); ??? Kohl, The Identification of Basis, 40 Tax L. Rev. 623, 654 (?) (1985) (offering no theoretical justification). Henry Simons himself ignored the tax treatment of borrowing, instead concentrating on such issues as unrealized gains and losses, gratuitous receipts, and in-kind compensation. E.R.A. Seligman, an influential writer in the early decades of the 20th century, appears to have bought into the accounting conceptualization of borrowing. See E.R.A. Seligman, The Income Tax 8-9 (2d ed. 1914); E.R.A. Seligman, Essays in Taxation 28-31, 102-06 (10th ed. 1925) (in context of discussion of property tax, stating that treating loans as lender assets but not negative assets of borrower is double taxation). The double-taxation argument would not, in a strictly formal sense, carry over to the income tax, where double taxation (in the current year) of the same dollars to two unrelated taxpayers is commonplace. However, current double taxation of borrowing might conceivably raise issues of policy that could influence tax system design. See ??? infra.

20 Joseph Sneed, Configurations of Gross Income 19, n. 16 (1967), advances two theories (without critiquing or analyzing them). One is the “equal value” theory. The other is that the taxpayer “invests in” (has a basis offset against) the borrowed money by virtue of the promise to repay the loan. The first theory ignores the critical role of basis, and the second is contrary to all other basis rules.

21 It has been suggested that the taxpayer has a (negative) basis in the repayment obligation equal to the tax-free value received on incurring the obligation (such as borrowed cash). It would follow that the disposition or satisfaction of the liability generates gain equal to the excess, if any, of such (negative) liability basis over the amount expended (having a positive basis) to dispose of the liability. Thus, if borrowed cash is repaid in full, here is no gain (or loss). See Amicus Brief of Professor Wayne G. Barnett in Comm’r v. Tufts, 461 U.S. 300 (1983). The negative basis theory is ingenious, but it is nowhere mentioned in the Code, the regulations, rulings, case law, or commentary. In any event, it is merely descriptive of a result that can be explained in other ways, such as the “inconsistent events” theory advanced by the majority opinion in the Tufts case.

22 Under double-entry book-keeping, the cash borrowed is debited to the Cash asset account, and the liability is credited to the Loan Payable liability (negative asset) account, rather than being credited to “income.” Similarly, the repayment of loan principal is credited to (reducing) the Cash asset account and is debited to (reducing) the Loan Payable account, rather than being debited to Expense. The payment of interest under both the income tax and accounting is treated as an “expense” (a current decrease in wealth), which is deducted (debited) against income of a business under business accounting rules and is deductible (or not deductible) for income tax purposes in accordance with section 163 of the Internal Revenue Code.

23 See Roswell Magill, Taxable Income 19-22 (rev. ed. 1945). Magill does not discuss the borrowing exclusion at all, but generally states that the legal concept of income is found in the assumptions employed by the usage of “income” in common speech as conditioned and modified by accounting practice, and in both ordinary speech and accounting practice borrowed funds are not viewed as income. However, the issue of whether tax theory should have a role in judicial interpretation of legislated text, which appears to be what Magill was addressing, is not the issue addressed in this article.
method of tax accounting can currently deduct the face amount of a future (expense) liability. The theory of accrual is that the taxpayer suffers a current decrease in wealth at the moment the obligation is incurred. These basic notions have not been challenged, and the only concern has been to shore up the boundaries surrounding the doctrinal category of “fixed and determinable liability” (talismanic of excludible borrowings and currently-accruable liabilities). Thus, the line is drawn between an “absolute” obligation to pay a determinable amount of money in the future (qualifying for offset against borrowing and accrual of deduction) and an obligation that is merely “non-ascertainable (in amount),” “contingent,” or “unreal” (which fails to qualify for current offset or deduction but is accounted for, if at all, only upon actual payment).

Upon reflection, it is apparent that the distinction between “fixed and determinable” future payment obligations and all other future payment obligations is itself a realization rule. Obligations that are contingent or without a fixed principal amount have no “face amount” that can be accrued to the present, whether as an offset against borrowing or as an expense deduction.

The liability-side rules are mirrored by those on the lender/creditor side. Just as borrowing is excluded, lending is not deductible, because the outflow of cash is (a “capital expenditure”) is exactly matched by the acquisition of an asset embodied in the repayment obligation of the borrower. And, just as principal repayments are not deductible by the borrower, so are they non-includible by lender on account of an equal basis offset. (Interest is treated semi-symettrically as an increase in wealth to the lender, included in income, and as a decrease in wealth to, but not necessarily deductible

24 See Comm’r v. Tufts, 461 U.S. 300 (1983) (reciting the rule that a borrower has no income because of the obligation to repay). No case has been found where the IRS has claimed that the proceeds of a true loan were income. In two major treatises on the income tax, the cases cited as authority for the proposition that borrowed money is not gross income all dealt with the issue of whether “advances” (against future amounts due as compensation, royalties, etc.) were income or excluded borrowings. See Friedich v. Comm’r, 925 F.2d 180 (7th Cir. 1991); Milenbach v. Comm’r, 106 T.C. 184 (1996); Matarese v. Comm’r, T.C. Memo.1975-184 (cited in Boris I. Bittker & Lawrence Lokken, 1 Federal Taxation of Income, Estates, and Gifts 6-2 note 2 (3d ed. 1999). In the early days of the income tax, accounting practices were taken seriously as giving content to “income” under the income tax. Thus, in Dilks v. Comm’r, 15 B.T.A. 1294 (1929), cash advances of contingent-on-profits compensation were held not to be income under a loan theory simply because the employer treated them as loans on its books.
25 See, e.g., United States v. Rochelle, 384 F.2d 748 (5th Cir. 1967), cert. denied, ??? (no borrowing where “borrower” had no subjective intent to repay).
26 There is an extensive body of doctrine holding that the receipt of rights that are contingent or forfeitable, and thereby raise difficult valuation issues, causes the received right to not be included in gross income (or to trigger current realization of gain or loss). See I.R.C.§ 83(a) (forfeitable property received as compensation); I.R.C. § 109 (lessee improvements received by lessor); Treas. Reg. § ??? (open transaction method of reporting contingent-payment sales); Burnet v. Logan, 283 U.S. 404 (1931) (contingent payment sale); M.E. Blatt Co. v. United States, 305 U.S. 267 (1938) (no income to lessor when lessee erects improvements); Comm’r v. Smith, 324 U.S. 177 (1945) (stock options before enactment of § 83).
27 The technical explanation is that the receipt of principal cash is fully offset by the lender’s “basis” in the repayment-obligation asset, which basis was created by the earlier capital expenditure (of making the loan). The function of basis in the income tax is to prevent the same dollars from being taxed more than once to the same taxpayer. In the present context, the lent amount has already been subject to tax (because of not having been deducted when the loan was made); thus, the receipt back of the “same” principal cash should not be taxed again.
by, the borrower.\textsuperscript{28} Tax symmetry also exists for accrual-method taxpayers: fixed and determinable rights to future payment are included in gross income at face. Moreover, such rights are included in “amount realized” (with respect to non-inventory property sales) by both cash-method and accrual-method taxpayers.\textsuperscript{29}

B. The Ability-To-Pay Critique of the Conventional Wisdom

The conventional wisdom regarding the current recognition of fixed and determinable (i.e., realized) future payment obligations is inconsistent with income tax theory based on the ATP norm.\textsuperscript{30} Briefly put, and issues of valuation aside, future payment rights generally represent current ability to pay of a taxpayer. Where such rights are truly illiquid, taxation of them should be – and is -deferred.\textsuperscript{31} In contrast, future payment obligations do not diminish current ability to pay. Thus, any attempt to argue that the borrowing exclusion is just the other side of the coin from the capitalization principle fails. In the capitalization scenario, properly conceived, the outlay is offset by the acquisition of an asset that is liquid or, if not, could have been liquid.\textsuperscript{32} In the borrowing scenario, the cash or asset acquired is not burdened with a corresponding diminution of ability to pay (except possibly if the liability is payable on demand). Before payment on an obligation is made, the taxpayer suffers no diminishment of assets or resources, nor does it lose control of assets or resources.

\textsuperscript{28} See I.R.C. § 163(h) (providing that personal-consumption interest, other than qualified residence interest, is not deductible).

\textsuperscript{29} Traditionally, such rights were included at fair market value by a cash-method seller and at face by an accrual-method seller. However, under current law the amount realized is generally the stated or “real” (present-value) principal under I.R.C. §§ 1273 or 1274. See Treas. Reg. § 1.1001(g) (1996).

\textsuperscript{30} Apparently, the only comment questioning the borrowing exclusion is William D. Popkin, The Taxation of Borrowing, 56 Ind. L.J. 43 (1980). Popkin nevertheless concludes that the tax treatment of borrowing is not “inherently objectionable” under an income tax (as a general rule) because debt-financed consumption is tax-advantaged relative to wage-financed consumption in the same way that wage-financed consumption is favored relative to savings-financed consumption. This point is a restatement of the fact that liabilities are treated as negative wealth, just as savings are viewed as positive wealth, and that the tax treatment of the former is the “mirror” of the tax treatment of the latter. However, this point is merely descriptive, and begs the question of whether liabilities should be treated as negative wealth. Popkin goes on to argue that not all liabilities should be treated as negative wealth under the income tax.

\textsuperscript{31} There are two lines of “gross income” cases that may have been wrongly decided (from an ATP perspective) under current law. One is the “economic benefit doctrine” applicable to cash-method taxpayers, see, e.g., Sproull v. Comm’r, 16 T.C. 244 (1951) (vested, but inalienable, future right to cash currently placed in trust or escrow is income), and the other is represented by United States v. Drescher, 179 F.2d 863 (2d Cir. 1950) (inalienable annuity received by employee is current income). Compare Comm’r v. LoBue, 351 U.S. 243 (1956) (stating that the granting of a stock option is not income where the option is non-alienable and subject to future contingencies). However, under I.R.C. § 83(a) and (c)(2) inalienability by itself does not bar the inclusion in income of vested property (other than stock options) received as compensation for services. Perhaps this rule is justified on the grounds that a nonalienability feature in the employment context would likely be self-imposed or the result of collusion to defer tax. See I.R.C. §§ 402(b) and 409A (confirming the economic benefit doctrine in the compensation context).

\textsuperscript{32} I am here suggesting that consistent adherence to the ATP principle would entail a thorough re-examination of capitalization doctrine, but that would be beyond the scope of this article. I am also suggesting that capitalization should continue to be the default principle simply to remove a strong tax incentive to invest in non-liquid assets.
The conventional equal-exchange justification of the borrowing exclusion—namely, that the receipt of cash (or an asset) is “offset” by an equal liability—violates income tax norms. Equal exchanges under an income tax are fully taxable except to the extent that the taxpayer (here, the borrower) has a basis in the thing transferred. However, the future payment obligation of a borrower or accrual-method taxpayer has no basis. Basis represents previously taxed dollars, but borrowing is neither a taxable event in itself, nor does it involve the present investment of previously-taxed dollars. Under general income tax norms, the even exchange of a zero-basis item for cash or an asset should produce gain.

The same logic applies to the accrual of expense deductions upon the incurring of a future payment obligation. It is settled that the giving of one’s own negotiable note cannot give rise to a current deduction. In Don E. Williams Co. v. Commissioner, 429 U.S. 569 (1977), the Supreme Court held that the giving of even a note payable upon the holder’s demand did not support a deduction for a cash method note-giver, because the taxpayer had not yet parted with any asset. Basis can exist in the future payment itself, because cash has a basis, but not in promising to pay money in the future.

It can be pointed out that accrual-method taxpayers can take expense deductions unsupported by basis. But one cannot deploy the existence of the accrual method as doctrine in an attempt to justify it in theory. Both the borrowing exclusion and the accrual-method deduction of future expense-payment obligations are based on the same idea, namely, that the fixing of future payment obligations entails a present decrease in wealth. But that is the very proposition that is being critiqued from the angle of ATP tax theory.

The no-basis difficulty cannot be finessed by such statements as “the basis comes from the future payments” or “the basis is borrowed from the lender,” as these statements merely re-package in a conclusory fashion the conventional wisdom, without citing any norms or logic. Cash always has a basis; hence, the basis in the repayment

---

33 Basis is a statutory requirement for depreciation and loss deductions, see I.R.C. §§ 165(b) and 167(c)(1), and is implicit in the case of expense deductions, which are usually made in U.S. cash. (The acquisition and spending of foreign cash by U.S. taxpayers raises additional issues that are not pertinent to the present discussion.) There is at least one instance in which the giving of appreciated property is treated as an expense in an amount equal to the property’s value, rather than its basis prior to the transfer. But in this case, gain is recognized on the transfer to the transferor, so that the expense deduction is made with after-tax dollars. See Reg. § 1.83-6(b)

34 Although it has been asserted, see note ???, that a liability has a (negative) “basis” equal to the amount received (borrowed), so that a satisfaction thereof by repayment of the same amount produces no gain or loss, this use of the term “basis” is simply an artificial construct to explain proper tax accounting for borrowing under the assumption that borrowing is excludible. In other words, the “theory” of “negative liability basis” begs the fundamental issue of whether borrowings should be excluded initially.

35 The classic example is the exchange of services (which have no basis) for wages.


37 See Calvin H. Johnson, ???

38 Basis in a non-cash asset derives from a capital expenditure, an actual income inclusion, or a proper income exclusion. The same is true for cash, except that even cash that was erroneously not reported as gross income has basis, first because there is no practical way to assign a zero basis to certain cash (cash
obligation cannot be derived from the anticipated future cash payments, which also must have a basis when paid, even though they are not be deductible. If the cash used to repay principal were to have a zero basis, then the satisfaction of an obligation with such cash would give rise to additional gain. In short, the same basis cannot do double duty. As to the claim that the basis is borrowed from the lender, basis never transfers from one taxpayer to another unless the parties are related and the Code specifically mandates such a result. Lead cannot be converted to gold by waving a magic wand.

At this point, advocates of the present system argue that, although basis is the appropriate way of accounting for assets, it is inappropriate in the case of liabilities. What is important is that a decrease in wealth is “realized” when the obligation is “fixed” and the amount (value) thereof is reasonably determinable. But however relevant the ability to value something might be under a Simons income tax, the current income tax is a realization income tax, and the ability to value something does not by itself constitute realization. Stated formally, “reasonable valuation” is a necessary, rather than sufficient condition, for realization. There still must be a current decrease in wealth, and, from the ATP point of view, there is not.

The existing tax treatment of borrowing is inconsistent with the cash-method of tax accounting followed by individual taxpayers (apart, perhaps, from the Schedule C). A cash-method taxpayer can ignore a future payment right on the income side (until the cash is received), while treating a borrowing transaction as a tax wash, which amounts to accrual-method treatment. Yet non-borrowing liabilities are not accrued (except, being fungible and often out of sight), and second (if more controversially) because the assignment of a zero basis to such cash would have the effect of removing the tax statute of limitations from affecting the finality of the erroneous exclusion.

---

39 See, e.g., Davis . The approach advanced by Professor Wayne Barnett, supra note , appears to be that the repayment obligation itself (rather than the borrowed cash itself) has a negative basis in the borrower’s hands that derives from the tax-free receipt of cash (or an asset). The negative basis is diminished to zero (with no gain or loss) if principal is fully repaid with non-deducted dollars, presumably on the theory that the basis in the non-deducted cash cancels out (absorbs) the negative basis in the liability. It follows that there is gain in the case where the liability is satisfied or cancelled at a time when there is negative basis remaining after all principal payments by the taxpayer. However, this analysis is just an elegant way of dressing up conventional doctrine in accounting garb. It doesn’t answer the question of whether borrowing should be excludible in the first place.

40 See I.R.C. § 1016(a) (reducing basis on account of prior depreciation and loss deductions with respect to the same asset).

41 Examples of transferred basis are found in §§ 362(a) (tax-free transfer to corporation), 723 (tax-free transfer to tax partnership), 1015 (gifts), and 1041 (transfers between spouses and certain ex-spouses).

42 Actually, only the principal obligation is “fixed.” The interest obligation becomes fixed (in the sense of “earned”) only with the passage of time. The issue of what exactly is being accrued under current law is discussed at infra.

43 As Henry Simons pointed out in 1938, the concept of “accrual” with respect to a future payment obligation appears to be incompatible with the concept of “realization” in its most common meaning, namely, conversion to cash. See Henry Simons, note , at 83. That is, if gain is realized upon receiving cash in a sale, then perhaps loss (or offset against borrowed funds) is only realized when cash is paid. This point was one among many that Simons deployed to ridicule (id. At 80-90) the realization principle as it was then formulated. Realization, however, is not confined to “conversion to cash” but includes in-kind increases and decreases in wealth that are “irreversible.” See text at note ...
perhaps, on Schedule C). Citing this inconsistency, however, doesn’t tell us anything
about the merits of the cash method and the accrual method as they pertain to future
payment rights and liabilities. That issue is dealt with in Part IV.

The tax treatment of borrowing might also be framed as an issue of “open” vs.
“closed” transaction. Current doctrine adopts the open-transaction method, because the
receipt of cash is not taken into income “pending” the future satisfaction of the repayment
obligation. The closed-transaction method derives from the “annual accounting”
principle under which events of tax significance are accounted on a year-to-year basis
without regard to what might happen in the future. The PID approach (prescribing
inclusion of borrowed funds and deduction of principal and interest when paid) clearly
accords with the annual-accounting approach. The annual accounting approach is
dominant in tax doctrine.\footnote{The seminal case is Burnet v. Sanford & Brooks Co., 282 U.S. 359 (1931),
holding that a short-term business venture was to be accounted for on an annual basis, rather than its conclusion.}
Instances of open-transaction treatment are (apart from borrowings) considered to be rare aberrations.\footnote{One such are is that of contingent-payment-consideration sales. See Burnet v. Logan, 283 U.S. 404
(1931), grudgingly acquiesced in by Treas. Reg. § 15A.453-1(d)(2)(iii). The open-transaction approach is a
relatively minor exception to the closed-transaction method prescribed by I.R.C. § 1001(a). The other
open-transaction doctrine is that pertaining to put and call options, where the consideration for the option is
held in suspense pending the exercise or lapse of the option. See Rev. Rul. 58-234, 1958-1 C.B. 279. This
document is critiqued in Calvin H. Johnson, ???.

By way of rebuttal to the last claim, the capitalization principle might be
advanced as a salient open-transaction doctrine. However, capitalization is based on the
theory that the expenditure does not result in any current diminution of wealth rather than
any assumptions about possible or likely future events. But the latter argument can also
be made with respect to future payment obligations, namely, that they subtract from
current wealth. That brings us back to main issue of whether the incurring of a future
payment obligation is a meaningful present decrease in wealth.\footnote{The claim that the PID betrays inconsistency in its treatment of assets and liabilities depends on the
merits of whether the two are distinguishable. Income tax doctrine as to future payment rights is basically
correct under an ATP norm, because notes, receivables, and other rights to future payment are “assets”
(property) that can be sold in commerce. Thus, these assets are “real” and represent current positive
wealth. Issues may exist around the edges, such as where the right is contingent or non-assignable or
where the value of the acquired right is less than its face amount. Nevertheless, the existence of these
issues does not undermine the basic principle that true acquisitions of positive wealth would be reckoned
currently under the PID. At the level of foundational theory, there is simply no such thing as an “income
tax” without capitalization (and its sidekick, basis). For what it is worth, non-symmetry between positive
and negative wealth rules is already a feature of tax accounting doctrine. First, after the enactment of
section 461(h), deduction accruals in advance of payments are much reduced relative to income accruals in
advance of receipts. Second, for cash method taxpayers the receipt of future payment rights will often
result in gross income, whereas the giving of one’s own payment obligation never supports a deduction.}
earliest time that the creditor has the right (and possibly the power) to cause the debtor’s existing stock of wealth (if any) to be reduced. A repayment obligation entails no sequestration of funds or removal of assets from the obligor’s control or use. Obligations are routinely satisfied by means other than sequestered wealth: possibilities include satisfaction out of current income, existing liquid assets, new borrowing, refinanced debt, third-party satisfaction of existing debt, discharge of existing debt in bankruptcy, and discharge of existing debt by creditors. And, of course, there is simply the possibility of not satisfying the obligation at all. Even a liability that is secured by specific property does not (normally) diminish the current use value of the property. Arguably, then, the decrease in wealth only occurs upon actual payment.

Any decrease in wealth attendant upon the incurring of a future payment obligation only exists from an “external” perspective. In the world of business and finance, fore-knowledge of future liabilities, both fixed and contingent, serves very useful purposes. But the government, in its revenue-raising capacity, has no interest in knowing about future liabilities. The government is not making investment or lending decisions. Rather, government operates on a year-to-year basis, and its revenue-raising function also operates on such a basis through the annual (tax) accounting system. In short, the concepts of income for business accounting and tax purposes differ.

Thus, the internal perspective would be determinative for fairness (ability-to-pay) purposes. The external perspective would be the relevant one for economics, because prices take into account positive and negative cash flows. However, individual taxpayers are not themselves market commodities. Whether the external perspective should nevertheless prevail is considered in Parts III and IV.

---

47 In cases where property secures a debt, there may be restrictions that preserve the lender’s security interest. However, restrictions of this type have been explicitly held by the Supreme Court not to result in any change of ownership for tax purposes. See Comm’r v. Brown, 380 U.S. 563 (1965) (known as the “Clay Brown” case). See also Frank Lyon Co. v. United States, 435 U.S. 561 (1978) (ownership not altered by fact that long-term lessee would probably be able to acquire the property at the end of the lease term at a bargain price); Fawick v. Comm’r, 436 F.2d 655 (6th Cir. 1971) (a patent transfer for royalties constitutes a sale notwithstanding restrictions on the licensee that serve to protect the licensor’s royalty interest).

48 It can be argued that the government is an investor because it can reduce its own net interest expense by paying down its debt. Moreover, government might assume that deficit spending will generate enough incremental private economic activity (and incremental taxes) to pay off its incremental future interest obligations. However, government can avoid borrowing money by raising taxes or lowering expenditures. Moreover, excess government expenditures over current revenues may not represent any kind of social investment. Nowhere in any government budget that I know of is there any calculation of expected incremental future taxes generated by deficit spending. Deficits may be created for wholly cynical reasons, such as to pay off favored constituencies or create fiscal pressure (to raise taxes or lower expenditures) on future generations and administrations. Finally, government (unlike any individual taxpayer) is (under prevailing Liberal political theory) merely a means, and not an end in itself.

49 See Henry Simons, supra n. 34, at 80-83; Thor Power Tool Co. v. Comm’r, 439 U.S. 522 (1979). The concept of income under business accounting is subsidiary to balance sheet accounting that shows net worth of the business, and net worth is basically historical costs less expired costs and less future payment obligations. The balance sheet paints a bleak picture of what might be available upon liquidation. Income under the current realization income tax, in contrast, is wholly independent of any balance sheet and is unconcerned with liquidation rights. Income under the income tax is dynamic, designed to show only increases and decreases in wealth of a taxpayer that occur during a given time period (the taxable year).
C. An Alternative Borrowing Paradigm: The PID Approach

The PID approach, which is the alternative to the conventional approach to future payment obligations, is described herein. Future payment obligations arise in various contexts. The categories discussed below appear to be the relevant ones under the PID approach.

1. Cash borrowing

The prototype transaction is cash borrowing. The “straight” PID approach to cash borrowing is straightforward. Borrowed cash would be current gross income. Principal repayments would be deducted as a matter of right when paid, because such payment is a non-consumption decrease in wealth, and disallowance of the deduction would cause the “same” dollars to be taxed twice to the same taxpayer (which would be contrary to basic tax norms). The borrowing inclusion would not create liquidity issues, because the borrowed amount can include additional cash with which to pay any resulting tax.

It needs to be pointed out that advocacy of the PID does not necessarily amount to advocacy of the cash method and rejection of accrual taxation across the board. The PID would exist within the income tax, which features the capitalization principle. Adoption of the PID approach to liabilities is consistent with accrual-method treatment of receivables. Moreover, although this is a peripheral issue, deductions under the PID could accrue when cash becomes due and payable, as opposed to the time of actual payment, the theory being that passing of the due date allows the lender or creditor to take action against the obligor. The PID could treat certain purchases of goods and services on credit as a deemed borrowing in an amount equal to the present value of the future obligation, coupled with a deemed payment in an equal amount for the goods and services, with any excess of the actual payment over the deemed payment being treated as interest. This is hardly cash-method treatment. Finally, as will be noted below, there are non-cash-flow methods of implementing the PID approach.

A basic issue is whether interest payments would be deducted as a matter of right or deducted only if the borrowed funds were used for a business or investment (as opposed to personal consumption) purpose. Since borrowed money would be included in gross income under the PID, it might be argued that all interest, including consumer interest, should be deductible as a cost of obtaining includible income. One might interject at this point by asserting that the resulting PID treatment of borrowing would be identical to the treatment of borrowing under a cash flow consumption tax (CT), where borrowed amounts are fully includible and both principal and interest are fully deductible as a matter of right. However, the better view is that interest is not a cost of obtaining borrowed funds, but of using them. The payment of interest is the cost of using

---

50 [cite Michael McIntyre ??] Under the existing income tax, loan origination costs of the borrower (other than “points” on home mortgages) are generally treated as capital expenditures that (if deductible) are amortized ratably over the loan period. Under a PID, it might be appropriate to treat such costs as a
another party’s funds over time. If interest were viewed as a cost of obtaining (rather than using) borrowed money, then current law has it all wrong: all interest, even business and investment interest, would be disallowed as a cost of obtaining tax-exempt borrowed money. In conclusion, the PID approach would disallow deductions for personal-use interest.

An alternative to the straight PID approach (hereinafter referred to as the “reverse PID approach”) would be to exclude the borrowing and disallow deductions for both principal and interest. The rationale of the reverse PID approach is that the borrowed amount is equal in present value terms to the deductible principal and interest (in the case of business and investment borrowing); hence, a “reversal” of straight PID treatment is the financial equivalent of straight PID treatment. However, this equivalence does not work for consumer borrowing, where interest would be disallowed.52

2. Liquid purchase-money debt

This category encompasses credit purchases of services, assets, and goods other than purchases of non-liquid assets dealt with in the next subsection.

Straight DIIT treatment is appropriate for the credit purchase of assets, services, and goods in all cases other than those where (1) PID treatment results in significant income “bunching” and (2) such bunching poses a liquidity problem. This last category is referred to as “non-liquid purchase money debt,” and is discussed in the next chapter. In the case of liquid purchase-money debt, straight PID treatment should be workable, keeping in mind that consumer interest would generally be nondeductible. Accounting shortcuts may be available. Thus, in the case of consumer credit, the credit card issuer or the merchant could be required to submit an information return setting forth the excess (or shortfall) of new debt principal over debt principal repayments. Since individual consumer purchases tend to be relatively small, it would not be difficult for a taxpayer to keep his or her PID-approach tax under control, or even to pay off consumer debt to reduce taxes.

3. Non-liquid purchase-money debt

A potential liquidity problem can arise if the item acquired with a substantial amount of purchase-money (two-party) debt is itself non-liquid. Certainly casual investors and individuals contemplating the purchase of homes and big-ticket consumer durables would strenuously object to the “straight” PID approach, with its large up-front income inclusions, and render it politically impossible of enactment.

reduction in includible borrowed principal, as well as a reduction in deductible principal repayments, if any.

51 The conclusion that interest is a cost of using money has generally been adopted in the current income tax, where interest is treated as a “period cost” (expense of using money). See I.R.C. § 163(a) (stating that interest is deducted as an expense when paid or accrued). Compare I.R.C. §§ 263A(f) and 266 (narrow exceptions to “expense” treatment). Moreover, the “character” of interest (as business, investment, or personal) is generally keyed to the way the borrowed funds are used. See Treas. Reg. § 1.163-8T(c).

52 Alternatives to straight PID treatment in the case of consumer loans are described in the text at note ???. 

13
The liquidity problem can be finessed in this situation by resorting to a present-value-equivalent approach that would avoid the up-front bunching of income. As noted above, in the case where interest would be deductible under the straight PID, an alternative (in present-value terms) would be the reverse PID approach of excluding the borrowed amount but to disallow the deductions for both principal and interest.\(^\text{53}\)

In the case of personal-use non-liquid purchase-money debt, straight PID treatment would entail an income inclusion for borrowing followed by a deduction for principal payments only. The reverse PID approach available to business and investment purchases would not be available here, because it would implicitly allow an interest deduction where such is not warranted. Here an option would be to include (in the year of purchase) only the excess of the borrowed principal over the present value of the principal-repayment obligation. However, this option would be capable of generating a large up-front tax. A more taxpayer-friendly option would be to: (1) wholly exclude the borrowing, (2) forego future principal and interest deductions, and (3) include imputed income annually at a rate equal to the interest paid.\(^\text{54}\) This option will be called the “interest-inclusion option.”

In cases where consumer interest is deductible (as is presently the case with “qualified residence interest”),\(^\text{55}\) the reverse PID approach can be used.

4. Obligations not related to cash borrowings or purchases

Items in this category include obligations to pay royalties, tort and worker-compensation liabilities, liabilities under self-insured medical plans, tax liabilities, interest, liabilities under warranties, and self-insured health plan costs. Although these business costs in many cases are indirectly related to the acquisition of goods and services, they are not part of their purchase per se. Hence, if this category is to be recognized at all, it is better to identify it by a bright-line test rather than under a facts-and-circumstances inquiry pursuant to one or more standards.

The major issue that arises in this context under the PID approach is whether business-expense obligations of this type should be treated like borrowings. If they are treated like borrowings, then (under the straight PID approach), the inclusion of the borrowed amount will offset the current accrual deduction of such amount, and the payment of principal and interest would be deductible. Interestingly, this treatment equates with cash-method treatment! In fact the straight PID approach manifests a cash-

\(^{\text{53}}\) The asset (securing the debt) could be deemed to acquire a basis immediately (as under current law), but such basis should, of course, be figured with reference to the “real principal” of the debt.

\(^{\text{54}}\) Excluding the borrowing is the equivalent of including it but allowing a deduction equal to the present value of interest and principal. Inclusion of the imputed interest would be a kind of “recapture” of the “erroneous” ex ante deduction of the interest. In cases where the consumer interest would be deducted (as is presently the case with qualified residence interest under I.R.C. § 163(h)(3)), the imputed interest income would “offset” the deductible interest, so that the interest component would simply be ignored (neither deducted nor included).

\(^{\text{55}}\) See I.R.C. § 163(h)(3).
method approach to liabilities that arise from borrowing (or purchase-money debt) transactions. The obvious implication is that liabilities that do not arise from borrowing should generally be treated according to the cash method.

5. Distinguishing principal from interest

The distinction between “principal” and “interest” has significance in two areas: (1) in cases where interest is not deductible, and (2) in cases (not involving borrowed cash) where it is necessary to determine the amount spent on an item (because the amount spent is either an expense deduction or creates basis). The identification of implicit interest (and of “real principal”) is a concern where there is no stated interest or the stated interest is unnaturally low. However, identification of implicit interest is not always worth the effort. For example, professionals often do not charge interest on services and merchants often use “no interest” promotions, but the tax stakes might be insignificant. Current law already has an elaborate set of provisions that deal with the issue of when implicit interest and real principal are to be identified. Basically, the “real principal” is the present value of all future payments specified under the contract.  

To illustrate, if a business purchases services to be performed in Year 1 for $100K, with payment due in 24 months with no stated interest, there would be a PID deemed borrowing income inclusion in Year 1 of $89K (assuming a discount rate of 6% per year), and $89K would be treated as the deductible cost of the services. The $100K would be deductible as principal and interest when paid. Note again that the transaction nets out to a $100,000 deduction in Year 3, which is identical to cash-method treatment.

Cash-method treatment avoids the problem of figuring out the real principal in the case of open-account debt, which is common in the business context, and other liabilities not specifying a payment schedule, because in these cases present-value analysis doesn’t really work. Indeed, if the debt principal cannot be ascertained, then accrual (at present value) would appear to violate the realization principle, and that is an independent reason (apart from the PID approach) to account for liabilities of this type on the cash method. This point is developed later on.

6. Summary of PID rules

The foregoing can be summarized in the following outline for easy reference:

A. Business and Investment Debt


---

56 Various Code provisions define “principal” as the present value of future principal and interest obligations (as opposed to the stated principal) in certain situations involving below-market loans. See I.R.C. §§ 471, 1272-1278, 7872.

57 Cf. Treas. Reg. § 1.1275-4(c)(4) (retroactive calculation of interest under contingent-payment sales contracts). Retroactive calculations are incompatible with the accrual-method approach.

58 See text following note ???.

---

B. Consumer Debt

1. Cash borrowing: straight PID treatment (but with no interest deduction).
2. Liquid purchase-money debt: straight PID treatment, figured on a “pool” basis.
3. Nonliquid purchase-money debt: interest-inclusion option (exclude the borrowing, do not deduct the interest and principal, and include in income an amount equal to the interest accrued).

II. Doctrinal and Practical Pros and Cons of the PID

This Part examines the doctrinal and practical pros and cons of moving to a PID approach. This Part can be skipped by the reader interested primarily in the economic efficiency issues posed by the PID approach.

A. Doctrinal Advantages and Disadvantages of Moving to a PID

This section discusses the doctrinal pros and cons of moving to a PID.

1. Debt doctrine problems

The tax treatment of debt that was borrowed from accounting has generated doctrinal problems that would vanish under the “straight” PID.

a. Untangling debt and property accounting

The borrowing-exclusion rule has resulted in the improper linkage of debt and asset accounting. The story begins with United States v. Kirby Lumber Co., 284 U.S. 1 (1931). There the taxpayer corporation issued bonds, essentially receiving (borrowing) cash in an amount equal to the face amount of the bonds. The borrowed cash was, of course, excluded according to the conventional accounting rationale. The bonds were subsequently repurchased by the taxpayer in the market for substantially less than par, thereby eliminating the entire liability. The Supreme Court held that this transaction produced gross income due to the fact that the taxpayer’s “net worth” (assets less liabilities) was increased, because the re-purchase reduced the taxpayer’s liabilities more than it reduced its assets.

Kirby Lumber incorrectly linked the tax treatment of liabilities not just to the taxpayer’s Cash account but also to its entire array of assets. By so doing, it begat the “insolvency rule,” under which debt relief was considered to be income only to the extent it increased the taxpayer’s overall net worth. The insolvency exception, which has been
codified,\textsuperscript{59} is just plain wrong, because unrepaid cash borrowing is an accession to wealth regardless of the taxpayer’s other assets.

The Court in \textit{Kirby Lumber} also fostered the notion that the tax treatment of purchase-money debt was linked to the tax treatment of the debt-financed property by citing and distinguishing, but not overruling, the earlier case of \textit{Bowers v. Kerbaugh-Empire Co.},\textsuperscript{60} in which the Court had held that debt-discharge income was negated by a loss in the debt-financed assets over the same period. Although the \textit{Kerbaugh-Empire} case is now considered to be a dead letter with respect to its debt-discharge holding,\textsuperscript{61} its spirit (of linking the tax treatment of purchase-money debt to that of the acquired asset) has been kept alive, perhaps inadvertently, by the Supreme Court’s decision in \textit{Crane v. Commissioner}, 331 U.S. 1 (1947), involving the issue of whether purchase-money-debt dollars should be included in the basis of the purchased asset.\textsuperscript{62} The general function of basis is to prevent the double taxation of the same investment dollars to the same taxpayer. In most cases, the basis of an asset represents previously-taxed dollars\textsuperscript{63} or (non-borrowed) permanently exempt dollars.\textsuperscript{64} Purchase-money debt poses a problem in that, although the borrowing is excluded from income, the exclusion is only “temporary” due to the fact that loan principal repayments will be made in the future with non-deductible dollars. In \textit{Crane}, the Court considered the possibility that purchase money debt would not be included in basis due to the debt-financed amount being “before tax.” The Court could have resolved this problem simply by noting that assets purchased entirely with borrowed cash acquire a basis equal to the cash purchase price,\textsuperscript{65} and that it made no sense to impose a different rule for purchase-money debt.\textsuperscript{66} The Court might also have pointed out that the borrowing exclusion is supposed to be “set right” by

\textsuperscript{59} See I.R.C. § 108(a)(1)(B).
\textsuperscript{60} 271 U.S. 170 (19??).
\textsuperscript{61} See Vukasovitch, v. Comm’r, 790 F.2d 1409 (9th Cir. 1986).
\textsuperscript{62} Actually, the property was acquired by inheritance, and the predecessor of I.R.C. § 1014 prescribed a basis equal to its value upon the decedent’s death, and this value exceeded the debt-financed amount. But the Court, troubled by the possibility of having a basis for inherited property far in excess of the basis for debt-financed property acquired by purchase, discussed the appropriate basis rule for property purchased in whole or in part with borrowed money.
\textsuperscript{63} Thus, assume that Bob uses $20,000 of taxable wages to buy stock of IBM Corporation. The cash outlay is a nondeductible capital expenditure creating a basis of $20,000. Basis is readily conceptualized as an “account” attached to an asset representing dollars previously taxed to Bob, and continuing to be “after tax” by reason of not being deducted on account of being a capital expenditure.
\textsuperscript{64} Including the tax-exempt dollars in basis preserves the integrity of the exclusion. Thus, if Bob’s $20,000 purchase price in the previous footnote came from an excludible damages recovery, failing to give Bob a basis of $20,000 in the stock would result in erasure of the exclusion upon a subsequent sale.
\textsuperscript{65} It is impractical to treat borrowed cash as itself having a zero basis, because cash is fungible. If cash did have a zero basis, then exchanging the cash for anything in the market would produce gain equal to the amount of the zero-basis cash. Treating borrowed cash in this fashion would be tantamount to treating the borrowing itself as gross income, but with delayed realization.
\textsuperscript{66} Where borrowing is directly linked to the purchase of property or services, the zero-basis cash problem described in the preceding footnote could be finessed by excluding the debt-financed amount from basis, in order to “preserve” the before-tax status of the borrowed cash. However, a less favorable rule for purchase-money debt than cash borrowings would distort the credit and asset markets for no good reason. Moreover, such a distinction would have generated line-drawing issues. For example, would credit card transactions be treated as separate borrowing-and-spending events or as debt-financed purchases? Finally, the logic of the rejected rule would cause re-thinking of the tax treatment of debt-financed expenses.
repaying principal with non-deductible dollars, not by increasing the gain (or decreasing the loss) when the asset is ultimately disposed of.

A contrary result in *Crane* would have hopelessly entangled debt accounting with asset accounting. Unfortunately, the Court’s opinion still (and unnecessarily) muddied the waters by stating that excluding borrowed money from basis would cause future loan principal payments to be added to the basis of the property, a result which would have fouled up the calculation of depreciation deductions. This statement at least implies that there is some connection between debt and asset accounting, specifically, that the basis in a debt-financed asset somehow derives from future principal payments on the debt. If such is indeed the case, then failure to repay the principal should result in a basis reduction. This way of thinking led to the “purchase price adjustment rule,” under which a discharge of all or a portion of seller-provided purchase-money principal debt, for whatever reason, reduces the asset’s basis rather than giving rise to income. In addition, it was long thought that even a reduction in third-party purchase-money debt would likewise reduce basis, although this notion has recently been put to rest. Apart from its theoretical incorrectness, a reduction-in-basis approach can result in such messy consequences as a negative basis and indefinite deferral of gain. In addition, the logic of the reduction-in-basis approach dictates that a discharge of debt incurred to pay an expense should be treated as a reduction in the expense, a rule that would pose theoretical and practical difficulties.

It is now understood that the better approach to the non-repayment-of-debt scenario in general is to reverse the prior borrowing exclusion. Thus, borrowing is viewed as being excluded “on the assumption” that loan principal will be fully repaid, and if the borrower escapes repayment then the un-repaid principal should be “taken back into” the borrower’s current income one way or another, without affecting basis. However, there remains a vestige of the intermingling of basis and asset accounting. Under *Crane*, the corollary of including debt in basis was to treat the relief from such debt (on a sale, disposition, or abandonment) as an “amount realized.” Thus, the “income” that is generated by purchase-money debt relief acquires its “character” (as capital gain, etc.) from the asset transaction rather than the debt transaction. This result was taken to an extreme in the Supreme Court’s 1983 *Tufts* decision where it was held that even that portion of a non-recourse liability (that the buyer took over) that exceeded the property’s fair market value was included in the seller’s amount realized. Although the overwhelming weight of opinion is that any excess of a non-recourse liability over the

---

67 See I.R.C. § 108(e)(5).
69 The case of Zarin v. Comm’r, 92 T.C. 1084 (1989) (reviewed), rev’d, 916 F.2d 110 (3d Cir. 1990), generating six separate opinions, can be characterized as such a case. That case involved casino-financed nondeductible gambling losses. One view of facts in this case is that the debt reduction ultimately involved a reduction in a nondeductible expense or loss. The opposing (and better) view is that it resulted in debt-cancellation income.
70 It is improper to include the income in the year of borrowing, because the exclusion of the borrowed amount was doctrinally proper in that year. Another reason for not going back to the year of borrowing is that the tax statute of limitations may well have run with respect to that year.
property’s value should have been treated as ordinary debt-discharge income,\textsuperscript{72} the \textit{Tufts} rule has been codified.\textsuperscript{73}

Under the straight PID approach, debt (including purchase-money debt) would be fully includible in the year incurred. Thus, borrowed money would be “after tax” and supportive of basis and expense deductions. In most cases, the failure to pay principal, for whatever reason, would simply have the effect of precluding any deduction for principal repayments.\textsuperscript{74} In the case of a disposition of property subject to a liability, the amount of the liability assumed would have to be treated (as under the \textit{Crane} doctrine) as a payment of cash by the buyer to the seller, who then uses the cash to satisfy the liability.\textsuperscript{75} The straight PID approach could also accommodate the purchase-price-adjustment rule, as well as its sibling, the poorly-named “disputed-debt doctrine.”\textsuperscript{76}

As has been noted above,\textsuperscript{77} enactment of a straight PID across-the-board is unlikely. Under the reverse PID approach and the interest-inclusion approach, the borrowing will not give rise to current income, and in those cases, the existing approaches would continue to operate.

\textbf{b. Contingent repayment obligations}

Given that current law excludes borrowed money, it is necessary to determine the limits of the exclusion by reference to some concept of what constitutes a “true” borrowing. This issue has long created problems for the Supreme Court. First the Supreme Court held in \textit{North American Oil Consol. v. Burnet}, 286 U.S. 417 (1932), that cash received after a trial court judgment for civil damages was to be included in current income as steps.

\textsuperscript{72} See Brief of Professor Wayne Barnett, supra note ???; Geier, ???; Seto, ???. Such excess is treated as actual or potential debt-discharge income in the case of a recourse debt. See Rev. Rul. 90-16, 1990-1 C.B. 12.

\textsuperscript{73} See I.R.C. § 7701(g).

\textsuperscript{74} It might be claimed that under the PID approach the borrowed money would generate a double tax benefit for the same dollars: (1) deduction or capitalization of the payment made with the borrowed cash, and (2) deduction of the repayment of borrowed principal. However, the repayment is made with different (subject-to-tax) dollars than the borrowed dollars, which are immediately spent.

\textsuperscript{75} Assume 100\% bank financing for an asset purchased by X for $100K, which rises in value to $160K and is sold for $60K cash with the buyer taking over the liability of $100K. Under the straight PID, the sales price would be treated as being $160K in cash, resulting in gain of $60K and a principal-payment deduction of $100K.

\textsuperscript{76} These doctrines only apply where the seller of goods or services sells to the buyer on credit. Here, as opposed to third-party lender situations, there is no external monitor of the value of the goods and services, meaning that the value of the goods or services (and the purchase money debt) may be initially overstated. If, due to an inherent defect in the goods or services purchased on credit, the credit purchaser subsequently negotiates the principal amount owed downward, it would not be irrational to treat the amount of the reduced debt as never having been borrowed and spent. In that case, if the initial borrowing were reported as income, the downward adjustment in the purchase price should generate a current deduction relating back to the “excess” borrowing. Where the borrowed amount was capitalized into the basis of an asset, any basis in the purchased asset would be reduced by the excess amount. If the credit-purchased item produced a deductible expense in an earlier year, the simplest solution would be to allow both the earlier borrowing inclusion and the deduction to stand, rather than requiring a current deduction and an offsetting inconsistent-events inclusion in the current year.

\textsuperscript{77} See text following note ??? supra.
income where the taxpayer had a “claim of right” to the funds, despite the possibility that the taxpayer might have to restore the funds if the judgment were reversed on appeal. Next, the Court held in *Commissioner v. Wilcox*, 327 U.S. 404 (1946), that embezzled funds were excluded on the ground that title to embezzled funds was “void,” thereby flunking the requirement that income had to obtained under a “claim of right” but satisfying the “borrowing” characteristic of an unconditional obligation to repay. Subsequently, the Supreme Court held in *Rutkin v. United States*, 343 U.S. 130 (1952), that funds acquired by extortion were gross income notwithstanding the fact that the victim had a right to restitution. 78 Not long thereafter, in *James v. United States*, 366 U.S. 213 (1961), the Court overruled Wilcox and held that an embezzler had gross income despite having an obligation to repay the funds under state law.

After this series of cases, it appeared that the borrowing exclusion did not apply where (a) the repayment obligation was contingent on future events or (b) the repayment obligation, although legally absolute, was contingent in the practical sense because the obligor did not evidence a “consensual obligation to repay.” But doctrine was muddled again as the result of the Supreme Court’s decision in *Commissioner v. Indianapolis Power & Light Co.*, 493 U.S. 203 (1990), holding that security deposits received by a utility company were to be treated as excludible borrowings notwithstanding the fact that repayment was contingent on the customer paying its bills. The Court treated the non-repayment scenario as a conflation of (a) a deemed repayment followed by (b) the deemed application of the deemed repaid amount to the payment of delinquent utility bills. Although this characterization of the transaction is plausible, the Court did not explain why it was superior to treating the deposits as receipts subject to contingent (in practice) repayment obligations.

The distinction between contingent and absolute obligations may be viewed as an application of the realization principle to repayment obligations. Contingencies render the repayment obligation impossible (or at least excessively difficult) to value. However, the valuation-difficulty issue involves a matter of degree rather than essential quality, and reasonable people can disagree over particular applications. For example, some contingent obligations (singly or as a group) may be capable of valuation on the basis of “track record” or statistics. Another example might be nonmarketable stock options, which may be easier to value than tax doctrine assumes. 79 In any event, the distinction between contingent and absolute repayment obligations would vanish under the straight PID, where all receipts subject to repayment obligations and possibilities would be included, with principal repayments being deducted. Nevertheless, the distinction between absolute and contingent liabilities would have to be maintained under the borrowing-exclusion variations of the PID approach.

c. Sham loans

78 Justice Burton, the lone dissenter in Wilcox, wrote the majority opinion. Four Justices dissented. The majority relied on (a) the rule that gains from illegal businesses were taxable, (b) the fact that the taxpayer had the control and use of the funds, and (c) the observation that tax administration would be rendered difficult if a taxpayer’s right to the income had to be determined.

79 Under Treas. Reg. § 1.83-7, stock options are rarely treated as having an ascertainable fair market value. Modern valuations techniques may provide a solution. See Myron Scholes & ???.

20
Another problem under traditional doctrine is posed by “sham purchase-money debt” such as existed in the well-known Estate of Franklin case, and which can be illustrated by a scenario in which a motel owner sells a motel, actually worth $1 million, for $10 million of consideration, consisting of a $60,000 down payment and a balloon-payment non-recourse note of $9,940,000 bearing $60,000 interest per year, with the seller leasing the premises back for $60,000 rent per year. Here the buyer claims a Crane basis of $10 million, which generates a net loss per year equal to the depreciation deduction (the rent income offsetting the interest payments). In Estate of Franklin, the Ninth Circuit held that the borrowing in such a scenario was a sham because both parties knew that it would never be repaid, given that (a) the principal amount far exceeded the value of the property and (b) the buyer/obligor was not personally liable on the debt. Since the debt was a sham, there was no Crane basis to depreciate, and the “interest” payments were deemed not to be deductible as interest.

The Estate of Franklin approach raises several problems. First (and foremost), the burden is on the IRS to discover schemes of this sort. Second, there is an issue as to whether two-party purchase-money debt is a sham in cases where the principal amount exceeds the value of the property by a slight or moderate amount. Third, if there is no loan, what is the proper characterization of the down payment, the interest payments, and the rental receipts?

These problems would be avoided under the straight PID approach, because the borrowed amounts would have been included in income, generating a huge tax liability relative to the value of the property. Transactions of this sort would never take place. Code section 465, dealing with non-recourse debt, could be repealed. However, the straight PID approach might not control this kind of case, but instead the reverse PID approach, under which the borrowing is excluded but neither principal nor interest is deducted. The nondeductibility of interest would dampen enthusiasm for this kind of transaction, but not as severely as would straight PID treatment.

d. Mismatched treatment of borrowing and investment

A major category of tax shelter is a debt-financed investment in which deductible interest is matched against income that is effectively exempt (in whole or in part) from tax. This phenomenon is discussed in Part III, but it is worth pointing out here that the

80 Estate of Franklin v. Commissioner, 544 F.2d 1045 (9th Cir.1976).
81 Congress first responded to this kind of scheme by enacting I.R.C. § 465, which limits losses to “amounts at risk.” However, a taxpayer might avoid scrutiny under this provision simply by asserting that the loan is a recourse loan. In 1986 Congress enacted I.R.C. § 469, which holds that passive activity losses are to be deferred regardless of the distinction between recourse and nonrecourse borrowings. However, not only is § 469 exceedingly complex, but also there is a broad exception for taxpayers in the real estate business. See I.R.C. § 469(c)(7).
82 Compare Lebowitz v. Comm’r, 917 F.2d 1314 (2d Cir. 1990) (stating test to be whether the debt “unreasonably exceeds” the value of the collateral), with Pleasant Summit Land Corp. v. Comm’r, 863 F.2d 263 (3d Cir. 1988), cert. denied, 493 U.S. 901 (1989) (holding that the debt is real to the extent of the value of the collateral).
complex and burdensome anti-tax-shelter rules of current law – which have only limited effectiveness in any event - could be repealed under a PID approach due to the less favorable tax treatment of borrowing.

2. Problems with the “all the events” test for accrual-method accounting

The tax doctrine concerning accrual of deductions for future liabilities was launched in the seminal case of United States vs. Anderson, 269 U.S. 422 (1925), a classic situation in which the government won the battle but lost the war. That case involved a deductible munitions tax liability “for” the year 1916 that came due in 1917, the year that the tax return, accompanied by payment, was filed. Because tax rates had risen considerably in 1917 to finance U.S. participation in World War I, the government (contrary to the usual posture taken in timing-of-deduction cases) sought to accrue the deduction in the earlier year, 1916, arguing (successfully) that the taxpayer, by accruing the taxes on its 1916 books, was bound thereby for tax purposes, because the then version of Code section 446 (enacted in 1916) provided that a taxpayer could make a return according to its books (kept according to business accounting principles). The Court in Anderson stated:

In a technical legal sense it may be argued that a tax does not accrue until it has been assessed and becomes due; but it is also true that in advance of the assessment of a tax, all the events may occur which fix the amount of the tax and determine the liability of the taxpayer to pay it.

In Anderson, the Court initially embraced business accounting practice as the test for accrual. Indeed, the 1916 statute allowing taxpayers to return income in accordance with their book-keeping practice was enacted in response to pressure from the (corporate) business community to allow taxpayers to follow business accounting methods for income tax purposes. Although the 1916 statute in question provided an exception that disallowed book accounting practices that “did not clearly reflect … income,” it was just assumed that generally accepted business accounting principles could not be considered to be unreflective of income. 


85 Neither of §§ 163(d) and 469 apply to business activity, and § 469 is not applicable to “real estate professionals” and most C corporations.

86 The taxpayer’s argument (by the legendary John W. Davis) before the Supreme Court relied principally on the language of the statutory provision allowing the taxes to be deducted “when paid.” See the report of Anderson at 70 L.Ed. 347, 348 (1926). The Court held that the statutory provision requiring the accrual method to be used (where the taxpayer kept its books on that basis) prevailed. The taxpayer did not seriously make an argument that the way it accounted for this particular item did not “clearly reflect income.” The government had an easy time showing that accrual in advance of the due date of a tax liability was proper accrual accounting. The all-events test seems to have been first set forth as such in the Solicitor General’s brief in Anderson. Ibid.
The “all events” test of *Anderson* soon took on an independent doctrinal life of its own in the tax world, as the Treasury regulations expressly allowed accrual accounting and defined “accrual” of both gross income and expense deduction items by reference to the all-events principle, with the added requirement that the amount of the income or deduction item be reasonably ascertainable. However, the all-events test is vague, resulting in much litigation and numerous hard-to-reconcile Supreme Court decisions. For example, the rationale of *Anderson* supports accruals for reserves for estimated future liabilities, expenses, and losses, because such future liabilities, etc., become inevitable (fixed) at an ascertainable date, the amount thereof can be reasonably ascertained, and reserves are sanctioned by business accounting. However, the Court rejected the use of reserves in *Brown v. Helvering*, 291 U.S. 193 (1934), reasoning that a reserve account is nothing more than an aggregation of liabilities (etc.) that, taken separately, are contingent. However, it is arguably unprincipled to allow some items (inventories and depreciation) to be accounted for on a mass basis while requiring other items to be accounted for on an item-by-item basis.

On a more mundane level, the issue of when a liability has become “fixed” has created difficult line-drawing issues. One distinction that has arisen is that between discretionary versus ministerial acts that fix liabilities. Another distinction is that between meaningful versus *de minimis* contingencies. On this front, it is settled that the likelihood of future non-payment due to financial hardship (or other constraints) falls into the *de minimis* category.

---

87 Contested items are considered to be contingent *per se*, even if the contest is frivolous. See Dixie Pine Products Co. v. Comm’r, 320 U.S. 516 (1944).

88 It can be said that the mass accounting for inventories and (during some historical periods) depreciation had a statutory basis. See, e.g., I.R.C. §§ 61(a)(2), 471, and 472 (inventories). However, the statutory provisions dealing with expenses and losses do not explicitly require that they be accounted for on an item-by-item basis, and I.R.C. § 446(a) allows accounting according to the taxpayer’s books unless, under I.R.C. § 446(b), such accounting “does not clearly reflect income.” Nevertheless, there are substantial prudential arguments against allowing reserves. For example, the statistics may be “cooked” or based on subjective factors, and accruals “at face” ignore the time value of money. These points are underlined by the fact that the statutory relaxation in 1954 of the tax rules barring reserves (and, the sibling of reserve accounting, deferral of prepaid income) had to be repealed retroactively the following year on account of the fact that the broad accounting discretion involved in determining reserves (and prepaid income deferrals) inevitably led to huge and unmanageable revenue loss. See Linda M. Beale, *Book-Tax Conformity and the Corporate Tax Shelter Debate: Assessing the Proposed Section 475 Mark-To-Market Safe Harbor*, 24 Va. Tax Rev. 301, 337-338 (2004).

89 Compare Continental Tie & Lumber Co. v. United States, 286 U.S. 290 (1932) (future ministerial act did not prevent present accrual), *with* United States v. General Dynamics Corp., 481 U.S. 239 (1987) (filing of employee claims under taxpayer’s self-insured health plan characterized as a discretionary type of action that rendered future claims contingent). *Anderson* itself was a “ministerial act” case, because the tax return for 1916 would have been filled out and submitted in 1917.


91 See Spring City Foundry Co. v. Comm’r, 292 U.S. 182 (1934); Georgia Book Depository, Inc. v. Comm’r, 1 T.C. 463 (1943). This rule, which emphasizes the “legal” over the “practical,” sits uneasily with *Anderson*, which reverses the emphasis. In any event, this is a good example of how a judicial formulation that may fit the facts of one case is unthinkingly carried over to another type of case.
That the Supreme Court itself was still dealing with issues under the all-events test more than 60 years after *Anderson* reveals a failure of legal doctrine. The failure is attributable to the fact that the application of the all-events test hinges mostly on matters of degree, not kind or quality, and is not grounded on any tax norm or concept. In hindsight, these doctrinal problems could have been avoided if the Supreme Court in *Anderson* had held that the tax liability could not accrue any earlier than the time it becomes due and payable. In fact, most accrued liabilities are routine accounts payable for goods and services provided to the taxpayer by other parties, and such trade accounts that are payable upon receiving the invoice from the provider, which is likely to occur shortly after the time the goods and services are provided. Moreover, it is hard to see how a liability can become due before it is earned by the provider of goods and services. An “earned and payable” rule would have (1) aligned tax accrual accounting with the ATP norm, (2) prevented premature accruals, and (3) been relatively easy to administer as a clearer standard of “realization.” In addition, as will be explained in Part III, it is prima facie the economically efficient rule.

3. Just what is it that is being accrued?

The question that captions this subsection refers to the issue of whether the “thing” that is offset against borrowed amounts (or is accrued as an expense deduction for a future liability) is (a) the present obligation only to make future principal payments or (b) the present obligation to pay principal and interest payments. The answer that is revealed by a perusal of history and doctrine is that offsets and accruals were “of” the principal amount only. The principal-only approach derives from the double-entry book-keeping method of business accounting, where interest is accounted for wholly apart from the principal. The same approach was adopted early on under the income tax with respect to borrowing, and has continued at the level of case-law doctrine to the present day. Thus, in *Commissioner v. Indianapolis Power & Light Co.*, 493 U.S. 203 (1990), the Supreme Court held that a non-interest-bearing “security deposit” was not gross

---

92 The *General Dynamics* case, supra note ???, was decided in 1987, and *Hughes Properties*, supra note ???, was decided in 1986.

93 An even better rule would be one that accrues a deduction at the time it begins to bear interest at a rate at least equal to the rate on short-term Treasury bills, because in that case the amount deducted would be close to the present value of the obligation to pay principal and interest. However, it is hard to see how such a rule could be fashioned by the courts.

94 Thus, the Court’s opinion in *Anderson* missed a beat when it stated that the tax liability in question wasn’t any different than the taxpayer’s other accrued liabilities. Most liabilities result from goods or services provided to the taxpayer.

95 An earned-and-payable rule might appear to be derived from the matching principle of business accounting. However, the resemblance is only coincidental. See generally, Deborah A. Geier, The Myth of the Matching Principle as a Tax Value, 15 Am. J. Tax Pol’y 17 (1998) (attacking the view that matching is a tax accounting norm).

96 Specifically, the asset-account Debit (to cash or property) is offset by an equal Credit to liability principal, the satisfaction of the liability is marked by a Credit to Cash and a Debit to Liability, and stated interest (if any) is accounted for separately with the passage of time as a series of Credits to Cash and offsetting Debits to Expense.

97 See *United States v. Kirby Lumber Co.*, 284 U.S. 1 (1931) (holding that debt-discharge income was to be computed only with respect to the principal portion of a debt obligation).
income solely because of what the Court characterized as an “absolute” obligation to repay only the principal. In the realm of liability accruals, where the liabilities often do not bear stated interest, permitted accruals were (and even after the enactment of section 461(h), still are) in an amount equal to the face amount of the principal liability.  

The principal-only approach to liability offsets and expense accruals has proved to be inadequate for tax purposes, primarily because of a growing awareness that the value of a future payment obligation will not equal the stated principal amount unless the unpaid obligation bears a market rate of interest. In contrast to business accounting, which avoids valuations, the tax concept of income is intensely dependent on valuation, because the crux of income measurement is “realized changes in wealth.” It was eventually realized by the tax community that (1) the negative wealth represented by a loan repayment obligation could in some cases differ from the stated principal amount and (2) the negative wealth represented by expense accruals could differ from the principal face amount. Although determinations of fair market value would be cumbersome as a general matter, in the case of future payment obligations there is a ready proxy, namely, the sum of the present discounted values of all principal and interest payments under the obligation. Tax avoidance occurs under traditional accrual-at-face doctrine in cases where the stated principal of a (re)payment obligation exceeds the “real principal” (the present value of future principal and interest payments). This overstatement of loss will occur in cases where (a) there is no provision for interest, (b) the stated interest rate is lower than the discount rate for risk-free obligations, and/or (c) there is (uncompensated) risk with regard to the timely collection of principal and interest payments as they become due.

The loss to the Treasury on account of overstated principal became recognized in the high-interest-rate climate of the 1970s and 1980s, but, because the courts were unwilling to adapt doctrine to the “new realities,” the Treasury sought and obtained legislation from Congress. The key move towards acknowledgement of present-value analysis in the tax law was the 1982 statutory recognition of original issue discount (OID), and the 1984 extension of the OID idea to identify the “real” principal (the

---

98 See ???.
99 See Dean v. Comm’r, ??? T.C. ??? (refusing to find imputed interest with respect to interest-free loan to employee). In Mooney Aircraft Inc. v. United States, 420 F.2d 400 (5th Cir. 1969), and Ford Motor Co. v. Comm’r, 71 F.3d 201 (6th Cir. 1995), it was held that long-term non-interest-bearing liabilities could not be accrued at face. In Mooney Aircraft, the deduction was deferred (until, presumably, it was paid), but the rationale was framed in terms of the matching principle, i.e., that the long time period destroyed any possible nexus between the (current) revenue and the related future cost. The Ford Motor case is one of the very few cases to have openly embraced time-value-of-money analysis, but the actual remedy imposed by the Commissioner was a quasi-cash-method remedy. Nevertheless, the Sixth Circuit appeared to reject the notion that the Commissioner is barred from imposing a present-value solution in cases falling outside of the Code provisions adopting such a solution. Accord Prabel v. Comm’r, 91 T.C. 1101 (1988) (reviewed), aff’d, 882 F.2d 820 (3d Cir. 1989) (disallowing Rule-of-78s method of computing interest).
100 I.R.C. § 1272 (accrual inclusion of OID with the passage of time) was added by the Tax Equity and Fiscal Responsibility Act of 1982, § 231(a). An example of an OID obligation is where $10,000 is borrowed in year 1, there is no stated interest, and in year 5 there is a stated principal repayment of, say, $15,000. Prior to the enactment of § 1272, the lender claimed that no income was realized until the $15,000 was received after 5 years, and that the resulting gain of $5,000 was capital gain under the predecessor of Code § 1271(a). Under Code section 1272, the lender is now treated as having lent a “real”
present value of all future payments) in other situations.\textsuperscript{101} For example, in section 7872 of the Code, below-market borrowings in certain situations (mostly involving related parties and parties, such as employers and employees, involved in an ongoing commercial relationship) are treated as producing an immediate transfer of wealth from the lender to the borrower in an amount equal to the excess of the lent amount over the present value of the repayment obligation.

Also in 1984, Congress enacted section 461(h) of the Code\textsuperscript{102} in response to criticism of the rule that accruals of liabilities were to be at their face principal amount. But instead of overthrowing that rule in favor of a rule that allowed a deduction equal to the present value of the obligation at its inception, Congress (on grounds of expediency) enacted the “economic performance” rule that postponed accrual at face to an event that (usually) occurred after the date the all-events test would be satisfied. Basically, the economic performance test is satisfied according to whichever of the following is appropriate: (a) the date services or goods were provided to the taxpayer, (b) the date goods or services are provided by the taxpayer,\textsuperscript{103} or (c) the date of actual payment. The economic-performance rule was widely viewed, even by its creators,\textsuperscript{104} as an unprincipled second-best solution. The first-best solution was considered to be accrual at the time the obligation becomes fixed of an amount equal to the present value of the obligation as a package (the “real” principal amount) followed by accrual of (deductible) interest, but ignoring the payment itself.\textsuperscript{105} Section 461(h) has essentially smothered Anderson as controlling authority with respect to the accrual-method deductions, and in some cases postpones the accrual event beyond billing to the actual payment of cash.

Section 461(h) aligns the tax treatment of expense liabilities more closely to the PID approach. As noted above, in some cases the taxpayer would be placed on the cash method. In other cases, the cost (whether expensed, added to Cost of Goods Sold, or capitalized) would be deemed to occur when the taxpayer receives goods and services.\textsuperscript{106} It is argued in Parts III and IV that section 461(h) did not go far enough.

\begin{footnotesize}
\begin{enumerate}
\item principal amount of $10,000, with the discount element of $5,000 being required to be included in gross income as “interest” (even to a cash-method taxpayer) as it accrues.
\item See I.R.C. §§ 483, 1271-1288, and 7872. Most of these provisions were added by the Tax Reform Act of 1984, §§ 41 and 172.
\item Tax Reform, Act of 1984, § 91.
\item This category refers mainly to such items as warranty obligations and required (environmental) clean-up costs, and the rule that these items can be deducted when the services are provided mostly (but not always) results in the item being deducted when paid. See Treas. Reg. § 1.461-4(d)(7) Examples (1) and (2).
\item For example, suppose an accrual-method business taxpayer X buys business services from Y, agreeing to pay $10,000 at the end of 5 years, without interest. Assuming a discount rate of 5% compounded annually, the currently-deductible present value of the $10,000 face-amount obligation would be $7,835. Deductible interest would be compounded on an annual basis at the same rate (5%) as was used to discount the loan to its present value. Thus, the accrued interest over the first 12-month period would be $392 (.05 x $7835), over the second 12-month period $411 [.05 x ($7835 + 392)], and so on. At the end of 5 years, total accrued deductions would equal $10,000, and (because total deductions would have aggregated $10,000), the actual payment of $10,000 would be ignored.
\item Payment (to another) for providing services to a third party at a later date should be deducted when paid, not when the services are provided, contrary to the result under Treas. Reg. § 1.461-4(d)(7) Example 1(ii). Thus, categories (2) and (3) under present law would be treated the same.
\end{enumerate}
\end{footnotesize}
B. Administrative Issues under the PID Approach

Financial institutions, retail credit sellers, and the residential real estate industry would all strenuously resist enactment of the PID approach on the ground that it would reduce demand for commercial loans, consumer goods, and residential real estate. It is hard to discern any constituency that would favor a move to the PID approach. Given the unlikelihood of any shift to the PID approach, it will not be necessary to devote much space to its potential administrative problems. Nevertheless, it is worth noting them briefly.

First, under a regime in which cash borrowings and certain credit purchases would give rise to current income (or its equivalent), one can anticipate that certain taxpayers would engage in various strategies to conceal borrowings, such as: (1) borrowing from relatives, friends, foreign taxpayers, and accommodation parties and simply not reporting the borrowing at all, (2) treating advances received from such persons as excludible gifts, and (3) using tax-exempt intermediaries as conduits to conceal the identity of the true borrower. Imposing a third-party reporting (and perhaps withholding) obligation on domestic commercial lenders, etc., may be feasible, but imposing the same on relatives, friends, and foreign taxpayers would be viewed as intrusive and impractical.

Second, as noted earlier, there would be political pressure on Congress to enact modifications of the straight PID approach to preclude bunching, or even to retain the current approach to certain borrowings and liabilities. The various resulting statutory distinctions would need to be worked out. Strategies would emerge to disguise or alter the true character of borrowings so as to move them into tax-favored categories.

Third, some taxpayers would be expected to attempt to deduct principal payments in cases where the borrowing had not been reported as gross income (or was excluded under the law prior to the changeover to a PID). The enforcement problem here is analogous to that of enforcing basis rules under current law.

Fourth, as noted earlier, there would be situations where it would be necessary to distinguish interest from principal.

---

107 There is a feature of present law that might appear to follow the PID approach, namely, the treatment of borrowings against annuities and pension plans as (potentially taxable) “distributions.” See I.R.C. § 72(e)(4)(A) and (p). However, this type of transaction is the equivalent of borrowing from yourself, and the non-repayment has no consequences apart from leaving the reduced account balance as it is.

108 See I.R.C. § 7701(l) (giving the Treasury authority to issue regulations relating to conduit financing arrangements).

109 The enforcement problem here is analogous to that occurring under present law where basis is claimed for an item received in-kind and erroneously excluded from income, except the PID scenario can be expected to occur more often than present-law scenarios.

Fifth, some taxpayers might attempt to avoid the borrowing inclusion under a sham-loan theory, citing cases decided under the current doctrinal regime. This should probably fail, as taxpayers are generally bound by the form of the transactions.

Sixth, the PID system could be used to create income and deductions that can be readily assigned to taxable years where they can do the most good (or suffer the least harm). For example, college students could borrow large sums to be taxed currently in low tax brackets but to be repaid in high-earnings years. This ability to temporally shift income may not be a wholly bad thing, however.

Seventh, along similar lines, borrowings could be used to alter the amount and timing of foreign source income for the purpose of increasing the foreign tax credit limitation.

Other issues would undoubtedly emerge. It is hard to “weigh” administrative issues under a new untested tax regime whose precise contours are not known in advance. Nevertheless, this section has shown that there are several doctrinal and administrative advantages of moving to a PID. Whether the PID “works” in the realm of economics and finance is the subject of the remainder of this article.

III. MINIMAL ECONOMIC EFFICIENCY

This Part deals with the issue of the tax treatment of borrowing and liabilities in relation to the minimum threshold of economic efficiency, namely, avoidance of rendering unprofitable, after tax, investments that are profitable before tax. The analysis herein will generally ignore the possible variations in PID implementation noted earlier.

A. Tax Treatment of Debt-Financed Investments and Consumption

Under the current income tax, a future repayment obligation of a borrower is treated in a “mirror” fashion to an investment. A positive investment outlay is a nondeducted capital expenditure, and investment returns in excess of the capital expenditure (memorialized as basis) are included in income. This set of results is often said to result in the “double taxation of income,” first by way of disallowance of the deduction for the capital outlay (the purchase price being the present value of all future

---

113 See note ?? supra.
114 If the borrowed amounts are invested, the interest would be deducted.
115 [cite Fennell and Stark ??]
116 Borrowing from a foreign lender would create foreign-source income under general income-sourcing principles, see I.R.C. § 861(a)(1) (interest sourced to nationality of payor), without generating any related foreign tax (assuming the country of the lender has not shifted to a PID). This problem could be dealt with by creating a separate “borrowing income basket” under I.R.C. § 904(d)(1).
returns, including income returns), and second as the income portion of the return is received (or accrued) over time.

Borrowing is a negative investment under current tax law, because the excluded borrowed amount (an exclusion being the same as an equal deduction offset against income) is the present value of all future principal and interest payments, and the interest is (again) deducted when paid (or accrued), unless (as a general matter) the borrowing is to finance consumption. Thus, in present-value terms the negative investor obtains a double deduction for (business and investment) interest, just as the positive investor is subject to double taxation of income.

The double-taxation-of-investment-income point is deployed in tax policy debates in an attempt to show that, under the income tax, cash-financed investments are treated worse than cash-financed consumption, because cash-financed consumption is taxed only once, whereas income that is invested results in double taxation in the sense described above. Debt-financed current consumption is also taxed only once, assuming that the interest paid is not deductible, because the current non-deducted consumption outlay (financed with excluded borrowing) is equal in present value terms to the future non-deducted interest and principal payments.

Under the current income tax, debt-financed investment is treated even better than cash- or debt-financed consumption. First the nondeducted investment is made with excluded money. Second, the nondeducted principal payments are made (analytically speaking) with cash returns that are excluded from income as basis recovery. Third, the investment income inclusion is offset by deductible interest payments. In short, the debt-financed investor is ultimately taxed, on a deferred basis, on only the excess (if any) of investment income over interest paid.

Under a cash-flow consumption tax (CT), investments are deductible and all (consumed) returns are fully included. Borrowing again embodies the mirror treatment: the borrowed cash is included, and both principal and interest are fully deductible as a matter of right, even in the case of consumer loans. Both results derive from the CT norm of temporal neutrality: just as the deductible investment is the present value of includible returns, so is the includible borrowing the present value of deductible payments. It follows that the result for debt-financed investments is essentially the same as under the current income tax, because what is taxed is only the excess (if any) of total (consumed) future returns over total payouts of principal and interest. It is also the case that debt-financed consumption is taxed the same as cash-financed consumption, both being taxed once (as under the income tax).

These results can be illustrated by the following pair of scenarios. In the first, non-borrowed money (“Cash”), scenario, the taxpayer earns $10K as cash wages in year

---

118 The included borrowing purchases nondeductible consumption. Although the principal and interest payments are deducted, the payments must be made with included cash from other sources. In the investment example, the amount invested equals the loan principal, so that returns that are recoveries of investment can be the source of funds to pay the loan principal.
1 (but not year 2), and the taxpayer has the option of spending the cash on current consumption ("Cd Cash") or investment ("Id Cash"). In the second, borrowed-money ("BM"), scenario, the taxpayer borrows $10K in year 1 (for one year at 6% annual interest) to currently consume ("Cd BM") or invest ("Id BM"). The taxpayer has includible wage income in year 1 of zero and in year 2 of $10.6K that is used to repay the loan principal and interest.\(^{119}\) The four situations resulting from these assumptions are deemed to be equivalent to each other in that the taxpayer has either $10K of year 1 consumption or its year 2 present-value equivalent. The results (total gross income reduced to present value ("PvGI") using a 6% discount rate) under the consumption tax (CT), current income tax (IT), and (in the borrowing scenario) the PID approach, are arrayed below.

**TABLE 1 – CASH- AND DEBT-FINANCED INVESTMENT AND CONSUMPTION UNDER THE INCOME AND CONSUMPTION TAXES**

<table>
<thead>
<tr>
<th>Financed</th>
<th>Tax Type</th>
<th>Year 1</th>
<th>Year 2</th>
<th>PvGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Cash</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cd Cash:</strong></td>
<td>CT</td>
<td>include $10K</td>
<td>-</td>
<td>+$10K</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>include $10K</td>
<td>-</td>
<td>+$10K</td>
</tr>
<tr>
<td><strong>Id Cash:</strong></td>
<td>CT</td>
<td>wash-out(^{120})</td>
<td>include $10.6K</td>
<td>+$10K</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>include $10K</td>
<td>include $0.6K</td>
<td>+$10.57K</td>
</tr>
<tr>
<td>(2) BM:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cd BM</strong></td>
<td>CT</td>
<td>include $10K</td>
<td>include/deduct $10.6K(^{121})</td>
<td>+$10K</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>no incl./no ded.</td>
<td>include $10.6K(^{122})</td>
<td>+$10K</td>
</tr>
<tr>
<td></td>
<td>PID</td>
<td>include $10K</td>
<td>deduct $10K of $10.6K GI</td>
<td>+$10.57K</td>
</tr>
<tr>
<td><strong>Id BM</strong></td>
<td>CT</td>
<td>wash-out</td>
<td>include $21.2K/deduct $10.6K(^{123})</td>
<td>+$10K</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>excl./no ded.</td>
<td>include $11.2K, deduct $0.6K(^{124})</td>
<td>+$10K</td>
</tr>
<tr>
<td></td>
<td>PID</td>
<td>include $10K</td>
<td>include $11.2K, deduct $10.6K(^{125})</td>
<td>+$10.57K</td>
</tr>
</tbody>
</table>

\(^{119}\) The present value of includible wages under both scenarios, using a 6% discount rate, is $10K, thereby creating equivalence between the two without assuming any “outside” transactions or investments. It is assumed uniformly that any taxes due in year 1 are paid out of external funds.\(^{120}\) The included wages are offset by an equal deduction for the $10K investment.\(^{121}\) See the previous footnote, except that here neither the principal nor the interest is deductible.\(^{122}\) Since the borrowed money is consumed in year 1, the loan principal and interest, both of which are deductible, are paid with included year 2 dollars.\(^{123}\) The $21.2K GI consists of $10.6K included wages and $10.6 included investment return, and the $10.6K deduction is for the payment of principal and interest.\(^{124}\) The $11.2K GI consists of $10.6K included wages and $0.6K included interest, and the deduction of $0.6K is for investment interest.\(^{125}\) Here the interest and principal are deductible.
The foregoing yields an observation the significance of which is often overlooked. That is, contrary to the general argument of CT advocates that the IT is systematically biased against investment, this is not the case with debt-financed investment, which is ultimately taxed as it would be under the consumption tax: under both, only the excess, if any, of the actual investment return over the interest paid is taxed. Stated more abstractly, the “double taxation” of investment income under the current IT is compensated for by the “double deduction” of investment interest expense. This point is important, since lenders generally require only a relatively small borrower equity for debt-financed investments, and, as a consequence, most non-speculative investments can be preponderantly financed by debt.

A second observation from the foregoing is that the tax treatment of debt-financed consumption under a PID would be placed on a par with that of cash-financed investment under the IT. In both there is double taxation (in present value terms) of the same dollars. Debt-financed consumption would be double-taxed under the PID in the sense that interest paid is taxed twice: first, as the excess of the included borrowed cash over the present value of the deductible principal repayments, and second as the actual interest payment (made with after-tax cash) is disallowed. Of course, a shift to the PID would not disturb the fact that, under the income tax, cash-financed consumption would continue to be given preferential treatment relative to cash-financed investment.

Based on the foregoing, it would appear that the overall bias of the income tax against investment – which is overstated to begin with - could be significantly reduced by a move to the PID approach. However, it also has to be acknowledged that debt-financed investment under a PID would be taxed worse than under the IT (or CT), because the investment would be “taxed twice” while the borrowing would lose the “double deduction” of interest (the included borrowing equals the present value of the deductible principal and interest).

The loss of the double deduction for (business and investment) interest under the PID approach is inherent in the straight PID approach and explicit in the “reverse” PID approach noted earlier, where borrowing is excluded but both principal and interest are not deducted. Thus, the PID approach can be imported into present law simply by disallowing deductions for business and investment interest! In the case of debt-financed consumption, where interest is already disallowed, the alternatives are (a) straight PID treatment or (b) the “interest-inclusion” alternative.

In sum, adoption of the PID would confer worse tax results for both debt-financed consumption and debt-financed investment than would occur under a CT, so that it is not clear that the PID would reduce, overall, the existing IT bias in favor of consumption.

B. The PID Bias Against Debt-Financed Investments

126 [cite Noel Cunningham, Shaviro]
127 See text at note ???.
128 See text at note ???.
The PID approach treats borrowing worse than under the across the board. An argument might be made that this might be a good move even in the case of investments, since cash-financed and debt-financed investment would be taxed alike, thereby removing the existing tax bias in favor of debt-financed (as opposed to equity-financed) investment.\(^{129}\) There might also be moral objections to borrowing, or indifference based on the idea that debt-financed investments do not increase national savings, because the positive wealth represented by the investment is offset by the negative wealth represented by the borrowing.

These arguments are not credible from a policy perspective. Over time debt-financed investment increases net social wealth on two levels. First, the lender profits from the lending, and the borrower (in a world without tax) expects that the net economic return from the investment will exceed the interest cost. Stated simply, business and investment borrowing allows the same cash to be invested twice with the aim, in each case, of producing a positive before-tax net profit. Thus, in a dynamic economy, investment-motivated borrowing improves social welfare and should not be thwarted by the tax system.

Moreover, the suggestion made by Table 1 that debt-financed investment would produce the same results as for cash-financed investments\(^^{130}\) turns out to be illusory, because debt-financed transactions entail an interest cost that is absent in cash-financed transactions. In fact, the PID would “prohibit” a large portion of worthwhile debt-financed investment. In the case of a cash-financed investment, any tax on the net income return will be less than the net income, and therefore the investment will continue to be profitable after tax. In contrast, debt-financed investments, unlike cash-financed investments, entail an interest cost. A move to the PID approach would eliminate the “double deduction” for interest costs, but the “double taxation” of investment income would remain. In terms of the assumptions set forth in Table 1 above, where the investment return exactly equals the interest cost, there would be an additional tax cost on top of the interest cost. Thus, many economically worthwhile debt-financed investments would never be made under a PID. This point is illustrated in Table 2 immediately below, using the same assumptions as Table 1 (invested borrowed money of $10K), but omitting the year 2 wage gross income of $10.6K, and positing various before-tax net investment returns.

| TABLE 2 – DEBT-FINANCED INVESTMENTS UNDER THE PID |
| (all numbers represent thousands of dollars) |
| Net Income |
| Invest. Ret. | Pre-tax Net | Net Tax\(^{131}\) | After Tax Net |
| Year 1 | Year 2 |

\(^{129}\) Cite debt vs. equity literature ???
\(^{130}\) In both situations, as well as that for debt-financed consumption, Table 1 shows an amount subject to tax of $10.57K.
\(^{131}\) In present-value terms (assuming a discount rate of 6%), and assuming a constant tax rate, the inclusion of $10K in Year 1 is fully offset by a deduction of $10.6K in year 2. Thus, the “net tax” is the product of the Year 2 Gross Return (e.g., $600 in Year 1 assuming a before-tax return rate of 6%) and the tax rate (assumed to be 35%).
<table>
<thead>
<tr>
<th>Yield (%)</th>
<th>Spread</th>
<th>Before-Tax Rate of Return</th>
<th>After-Tax Rate of Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>6%</td>
<td>+10</td>
<td>+0.6 - 10.6</td>
<td>0</td>
</tr>
<tr>
<td>6.5%</td>
<td>+10</td>
<td>+0.65 - 10.6</td>
<td>50</td>
</tr>
<tr>
<td>7%</td>
<td>+10</td>
<td>+0.7 - 10.6</td>
<td>100</td>
</tr>
<tr>
<td>7.5%</td>
<td>+10</td>
<td>+0.75 - 10.6</td>
<td>150</td>
</tr>
<tr>
<td>8%</td>
<td>+10</td>
<td>+0.8 - 10.6</td>
<td>200</td>
</tr>
<tr>
<td>8.5%</td>
<td>+10</td>
<td>+0.85 - 10.6</td>
<td>250</td>
</tr>
<tr>
<td>9%</td>
<td>+10</td>
<td>+0.9 - 10.6</td>
<td>300</td>
</tr>
<tr>
<td>9.5%</td>
<td>+10</td>
<td>+0.95 - 10.6</td>
<td>350</td>
</tr>
<tr>
<td>10%</td>
<td>+10</td>
<td>+0.1 - 10.6</td>
<td>400</td>
</tr>
</tbody>
</table>

Under the assumptions used in Table 2, the before-tax rate of return would need to be more than 150% of the interest rate for the investment to break even after taxes. Thus, on its face the PID is violates the minimal threshold of economic efficiency.

The inefficiency inherent in the PID approach assumes that the before-tax income yield is fully taxed. In fact, the current IT is riddled with exclusions and their functional equivalents, including: (1) full exclusions on state and local bonds,132 returns on Roth IRAs,133 and gains on personal residences,134 (2) partial exclusion-equivalents for net capital gains135 and natural resources income,136 (3) income-deferral provisions relating to unrealized net gains, installment gains,137 like-kind exchanges,138 exchanges pursuant to entity restructurings,139 and future payment rights acquired by cash method taxpayers,140 and (4) full or partial expensing of assorted capital expenditures.141 A few Code provisions disallow expenses relating to tax-exempt income,142 but these provisions are construed narrowly to refer to wholly exempt income under a Code-mandated exclusion, meaning that income deferrals, deduction accelerations, and capital gains rates are not considered to produce tax-exempt income under these expense-disallowance rules. There are also complex sets of anti-sheltering rules for deferring net tax losses (or deductions) in situations where taxable income is expected to be shifted forward by reason of income-deferral and/or deduction-acceleration rules,143 but the broadest of these rules is limited to certain passive investments.144 Thus, opportunities abound under the

---

132 See I.R.C. § 103(a).
133 See I.R.C. § 402A(d).
134 See I.R.C. § 121. In addition, modern-day tax shelters are often structured so as to deflect taxable income (but not economic income) to tax-exempt persons. (cite article ???)
135 See I.R.C. § 1(h) (lower rates).
136 See I.R.C. § 613 (percentage depletion deduction, which operates in the manner of a partial exclusion).
137 See I.R.C. § 453.
138 See I.R.C. § 1031.
139 See I.R.C. §§ 351, 354, 361.
140 See Treas. Reg. § 1.446-1(c)(1)(i).
141 See, e.g., I.R.C. §§ 168, 174, 179, and 263(c); Treas. Regs. ??? (various expensing and accelerated depreciation provisions).
142 See I.R.C. 264(a), 265(a).
143 See I.R.C. §§ 163(d) (investment interest in excess of net investment income), 280A (rental use of home), 465 (non-at-risk activities).
144 See § 469 (also not applying to C corporations and real estate investments by real estate professionals).
income tax to combine the borrowing exclusion with tax-favored investment, thereby combining the “double deduction” of interest expense with something less than the “double taxation” of investment income. This synergy produces an overall tax result that can be better than that obtainable under the consumption tax standing alone, which (in present value terms) is a zero tax rate. In other words, combining the income tax treatment of debt with a full or partial consumption tax treatment of income produces a negative tax (government subsidy)\textsuperscript{145} that diverts resources to non-economic uses.\textsuperscript{146}

Under the PID, the same transaction as analyzed in Table 2, but with full exemption of economic returns, would be treated as follows: Year 1 net inclusion of $10,000, coupled with a Year 2 result of no gross income and deductions of $10,600. In present-value terms (at a discount rate of 6%), the economic wash would be matched by a tax wash. This (correct) result follows from the fact, noted above, that the PID treatment of debt is the functional equivalent of disallowing the interest deduction under a borrowing-exclusion regime. In other words, the PID approach operates in a manner that is the equivalent of Code section 265. Thus, assuming a full exemption for income (or its equivalent), PID treatment of debt would assure a result that is no worse than “straight” exemption (CT treatment) of investment returns. However, the PID approach will not produce the “correct” result if any of the gross return is taxed. Thus, referring to the first line of Table 2 (gross return and interest cost both being $600), if only one sixth of the gross return ($100) is taxed, there is still a negative after-tax return on an economic wash. Nevertheless, the higher the exemption on the gross return (or its equivalent), the less will the PID render unattractive a positive net investment. To generalize, the PID treatment of debt is “more” efficient than the IT treatment of debt where the effective income exclusion is greater than 50%. Thus, under the current corrupted income tax, it might be worth “trading in” (repealing) such anti-tax-arbitrage provisions as Code sections 163(d), 264, 265, 465, and 469, and perhaps the AMT, in return for enacting the PID treatment of debt. On the other hand, such a move could be seen as a surrender in the long-fought war to preserve the integrity of the income taxation of investments.\textsuperscript{147}

\textsuperscript{145} For example, assume a situation where $10,000 is borrowed at a 6% interest rate to purchase an investment that generates interest income of 6%, both over one year, so that economically the entire transaction is a washout. Assume also that the interest income is tax exempt. Under the current IT, there is a Year 1 tax wash (borrowing exclusion coupled with non-deducted capital expenditure), and in year 2 a tax loss of $600 (deducted interest coupled with excluded income), producing an after tax profit ($600 times the marginal tax rate) on the economic washout.

\textsuperscript{146} See Calvin H. Johnson, \textit{Is an Interest Deduction Inevitable?}, 6 Va. Tax Rev. 123, 128 (1986). The contrary argument that interest should be deducted against tax-exempt income is that otherwise the tax preference on the income side will disappear if the investment is debt-financed. See George Mundstock, \textit{Accelerated Depreciation and the Interest Deduction: Can Two Rights Really Make a Wrong?}, 29 Tax Notes 1253 (Dec. 23, 1985). However, no one is arguing that courts should disallow interest deductions in order to invalidate Congressionally-sanctioned tax shelters. The argument against the interest deduction against tax-exempt income is a normative one based on economic efficiency. Debt-financed investment is efficient where it produces a positive before-tax net return, not a zero (or negative) net return. Thus, there is a good reason to prefer cash-financed tax-favored investments over debt-financed tax-favored investments: the latter (can) decrease net social wealth.

\textsuperscript{147} See ??? Bankman (suggesting that that the war to tax investment returns is a losing one). The PID treatment of borrowing is the same as the consumption tax treatment of borrowing in the case of business and investment activity. Consistent consumption tax treatment of both investments and debt would avoid the mismatch of interest deductions against tax-favored income. Such a tax system might be unattractive to
C. Debt-Financed Current Business and Investment Expenses

In the previous section it was assumed that cash borrowing was “invested,” and it was concluded that PID treatment of borrowing would drive out good investments (investments that would be profitable before tax) if the income were taxed but would drive out bad investments if the income were not taxed (as where the investment, inherently a capital expenditure, is allowed to be expensed). PID treatment is also correct where the expensed item is “really” an expense (rather than a capital expenditure), meaning that the revenue “caused” by the expense occurs in the same taxable year as the expense. As an aside, borrowing is not really necessary here unless the revenue is not reaped until after the expense is paid. Assume the following scenario: (1) $100K is borrowed on April 1 to pay a $100K labor cost on April 1, (2) the revenue yield from this cost is $102K, reaped 120 days later, and (3) the principal and accrued interest (totalling $102K) is immediately paid off. Under the PID, the taxpayer includes a total of $202K and deducts exactly the same amount, so that the economic wash is correctly reflected by a tax wash.

It happens that the current tax treatment of borrowing also produces a correct result here: the $102K included amount is matched by aggregate deductions of $102K. Thus, it cannot be claimed that the PID is superior to the IT in the “true expensing” situation. Nevertheless, it is also the case that the current income tax doctrine makes only a feeble attempt to identify those true expenses that ought to be deducted currently. In addition to the numerous statutory expensing rules already alluded to,148 various Treasury Regulations allow wholesale expensing for what are really capital expenditures.149 In other areas, capitalization is simply not enforced.150 Since the expensing of capital expenditures negates the “double” taxation of the income generated by such expenditures, PID treatment of the related borrowing would be superior to IT treatment with respect to the entire class of items actually claimed as expenses.

At this point, the prospect is raised that, at least on efficiency grounds, it would be best to install a two-track system, namely, PID treatment of borrowing to pay deductible expenses and IT treatment of borrowing to fund business and investment capital expenditures. The problem with a dual-track system for debt is that, since borrowed

---

148 See note ??? supra.
149 See, e.g., Treas. Reg. § 1.162-4 (“repairs”), -5 (certain costs of acquiring information and skills), -6 (costs of carrying on a profession), -12 (farming costs), - 20 (a)(2) (“goodwill” and other advertising); 1.263(a)-4 (expensing allowed for certain intangible-creation costs).
150 For example, internal law firm costs are often capital expenditures, since the fees won’t be collected until the future. In the case of contingent-fee litigation, the time lag may be quite significant. The IRS has apparently never attempted to enforce capitalization in this area. Another profession in a similar posture is architecture and design.
money is fungible, an allocation would be difficult.¹⁵¹ A rule that allocated business debt according to what assets were used as security for the debt would heavily stack the deck in favor of characterizing debt as investment-related (as opposed to expense-related), because assets can always be found to secure debt. A rule that characterized debt according to how the borrowed cash is actually spent would be almost impossible to administer in a business or investment context. A rule that made the allocation according to the ratio of total current expenses to total current capital expenditures would not be impossible,¹⁵² but it would be resisted as too complex.¹⁵³ Any allocation rule would invite manipulation and subterfuge, as taxpayers would clearly prefer debt to fall within the IT regime. Thus, the problem of allocation renders a hybrid IT/PID system for business and investment cash borrowing unlikely, at least for third-party loans.

In contrast to cash borrowing from third parties, two-party debt is necessarily “earmarked” to the goods or services purchased on credit, and therefore the allocation task mandated by a two-track system could be carried out. However, an incentive would be thereby created to avoid two-party financing of expensed capital expenditures.

D. Business and Investment Liabilities Not Arising from Financing Transactions

In an earlier section it was suggested that liabilities for future expenses (such as for torts, fines, and taxes) that do not arise from the borrowing of cash or the purchase of goods or services be deducted under the cash method. Under the PID approach, the explanation advanced earlier was that any deemed includible borrowing is offset by the accrued expense deduction, and the only deduction that stands is that for deductible interest and principal.¹⁵⁴

Considerations of minimal economic efficiency support this view. Indeed, the argument for cash-method treatment is valid under a conventional income tax as well. The argument is that there is no debt-financed investment in these instances. There is no “debt” because the obligee (if capable of identification) has not parted with money, goods, or services that would justify the charging of interest. There is no investment on account of the fact that the taxpayer has not yet received anything of value. Even if value has been received indirectly,¹⁵⁵ it would not have been capitalized as a cost. Thus, there would be no double taxation of investment that needs to be balanced by the double deduction of interest expense.

¹⁵¹ See Daniel I. Halperin, Interest in Disguise: Taxing the “Time Value of Money,” 95 Yale L.J. 506, 528-31 (1986) (discussion issue of whether nuclear decommissioning costs should be viewed as a current investment to produce current income or a future expense).
¹⁵² [cite allocation of interest rules under sec. 864(e).] An issue would be whether “current capital expenditures” would be net of capital recoveries for the year.
¹⁵³ One aspect of the complexity would be the seeming necessity for retrospective re-allocations attendant upon disputes over capitalization, deductions, and perhaps capital recovery items.
¹⁵⁴ See text at note ???.
¹⁵⁵ Future tort, tax, and fine liabilities might be considered a by-product of current business activity.
It might be argued that the taxpayer (if the present value of the liability is ascertainable) might set aside an investment fund that would provide the cash to satisfy the future liability, in which case such investment could be considered to be effectively debt-financed. However, there are other, more plausible sources for paying such future costs, namely: (1) future revenues, (2) insurance, and (3) future borrowings, not to mention the alternative possibility of non-payment. Any such future borrowing would be in the nature of a debt-financed expense payment, which is not inefficiently treated under a cash-method (PID) approach. In short, one should not postulate a tax-burdened transaction when a non-tax-burdened alternative is readily available.

Treatng the future liability as a deemed borrowing would actually operate as an unjustified tax shelter! Assume that the taxpayer purchases inventory for $80x and immediately sells it for $100x, but that it has future warranty obligations that have an ascertainable negative present value of $20x (discounted from a future value of $30x), so that the transaction as a whole is an economic wash in present-value terms. Suppose the warranty aspect of the transaction is treated as a current borrowing of $20x that will accrue $10x interest by the time of actual payment. Under an income tax in which the taxpayer is treated as borrowing $20x in the year of sale to pay a current liability, the taxpayer would be allowed to deduct the $20x currently (as an expense) and the $10x in the future (as interest), resulting in an overall tax loss of $10x. This result is produced by a double deduction of interest unmatched by any (necessary) double taxation of income. The correct result under both the PID and the income tax is to include gain of $20x in the year of sale and deduct $30x in the year of warranty satisfaction (or $20x in the year of sale).

Thus, the deferral-of-deduction-until-payment accrual rules found in Code section 461(h)(2) and the regulations for liabilities of this type appears to be correct in principle, and is not (as is commonly asserted) a second-best or imperfect solution. This claim is examined again in Part IV.

E. Debt-Financed Consumption

This section considers the implications of PID treatment of debt-financed current consumption and consumer assets. Recall that “personal interest” would be disallowed under the PID. Straight PID treatment would entail inclusion of the borrowed cash and deduction of only the principal repayments. Thus, the loan transaction would be subject to a tax (on the nondeductible interest payments) over and above the tax on the debt-financed consumption itself. This phenomenon can be called the “double tax on debt-financed consumption,” which happens to the exact analogue of “double taxation of

---

156 See I.R.C. § 461(h)(2)(c) (tort and worker compensation liabilities); Treas. Reg. § 1.461-4(d)(7)(Exs. 1 & 2) (warranty costs) and (g) (other future payments other than interest and property taxes).

157 See note ???.

158 See text at note ???.

37
(cash-financed) investment income" under the income tax. This result compares unfavorably with the result for cash-financed consumption, which is taxed once.

The effect of straight DIT treatment can be achieved by the alternative technique of ignoring the loan but including in income the interest owed as it accrues (or is paid). The existence of this alternative “interest-inclusion” method demonstrates that PID treatment of debt-financed consumption assets would have the indirect effect of taxing the imputed income from consumer assets by proxy! The prevailing (if not unanimous) view in academia is that net imputed income should be taxed within the income tax in order to prevent tax discrimination in favor of consumer assets relative to investment assets. The PID offers a round-about way of accomplishing this aim.

The foregoing suggests the viability of adopting the PID approach only for debt-financed consumption with the aim of removing, in significant part, the alleged tax bias of the income tax in favor of consumption over savings. Debt-financed consumption would appear to be more at the margin of the investment vs. consumption decision, since the fairly well-off have both the option to save or consume and the ability to borrow significant amounts of money.

It could be argued that debt-financed current consumption, despite the absence of imputed income, would be taxed the same way under the PID as debt-financed consumer assets. For most individuals, debt-financed current consumption would usually show up in the tax base as an increase in credit card principal balances at the end of the current year relative to such balances at the end of the previous year. Thus, the consumer who buys current consumption on credit can avoid worse treatment (relative to current law), simply by avoiding increases in such balances. Significant increases in consumption-related debt principal balances would likely result from precisely the purchase of consumer durables, vehicles, and homes, i.e., consumer assets that produce imputed income.

The analysis so far has assumed that consumer interest is nondeductible, but there are exceptions under current law, most notably for qualified residence interest, which is considered to be a political sacred cow. Under the current income tax, the

---

159 The PID tax on consumption borrowing equates with the IT tax on the nondeductible capital expenditure, and the PID tax on nondeducted interest payments equates with the IT inclusion of interest (income) returns.

160 See ??? But see Thomas Chancellor, ???.

161 The same technique could be used within a consumption tax. There, the ideal treatment of consumer assets is to treat them the same as investment assets, but to include net imputed income in the tax base. But taxing net imputed income directly has been considered impractical. Imputed income can be taxed indirectly (in the case of debt-financed consumer assets) as follows: include the borrowed cash, deduct the purchase price, deduct the loan principal payments, and disallow interest deductions. However, this approach seems inferior to allowing the interest deduction and disallowing a deduction for the purchase price, which amounts to ex ante taxation of the net consumption value reduced to present value.

162 See I.R.C. § 163(h)(3). Other categories of deductible consumption interest (arguably) include qualified educational interest (I.R.C. § 221) and interest (to the extent allowable) with respect to vacation homes, home offices, and hobby activities (see I.R.C. §§ 183(c) and 280A(c)).

163 [cite ???] Bush’s refusal to consider adopting his own Tax Reform Panel’s proposal.
deductibility of personal interest creates a tax shelter effect. Under the PID approach, allowing the interest deduction only eliminates the tax on the borrowing transaction as such. Thus, retaining a mortgage interest deduction under a PID would seem less objectionable than it is under current law.

I take no position on the issue of whether adoption of the PID move would be good economic policy. Preliminarily, imposing a higher tax cost on consumer borrowing might have an overall dampening effect on the economy, unless the resulting expanded tax base resulted in lower rates (which could have the opposite effect).\textsuperscript{164} Also, there could be effects not only with respect to the choice between consumption and savings,\textsuperscript{165} but also on the financial industry, where the substitution effect (reducing the demand for loans) would to some degree be offset by the income effect (increasing amounts borrowed per loan so as to cover increased tax liabilities). These issues are, however, beyond the scope of this article. From the technical point of view, the major issue would be whether the distinction between personal interest and business (or investment) interest could be maintained in this context, but the distinction already exists under current law, where it appears to have been maintained with reasonable integrity.\textsuperscript{166}

To sum up, the existing tax treatment of debt is preferable to the PID approach in cases where investments financed by such debt yield income that is more than 50% effectively taxed. However, the PID approach is acceptable for debt-financed expenses. Finally, PID treatment of debt-financed consumption might be considered as a way lessening the alleged income tax bias against savings.

IV. DOES ECONOMIC EFFICIENCY MANDATE ECONOMIC ACCRUAL OF ALL LIABILITIES?

An oft-advanced view is that all liabilities should be treated like cash borrowings are treated under current law.\textsuperscript{167} The borrowing model is really an original issue discount (OID) model in cases where there is no stated interest or below-market stated interest. Under the OID model, there would be (1) an immediate accrual of an expense deduction in an amount equal to (present) value of the future (principal amount) payment obligation and (2) identification (and deduction) of implicit interest as it accrues.

\textsuperscript{164} It appears to be common wisdom that consumer spending, much of which is debt-financed, drives the economy. ???
\textsuperscript{165} ???

\textsuperscript{166} This statement is hard to prove. Nevertheless, this issue is not specifically mentioned on any list of pressing IRS administrative problems, nor has any conspicuous literature on this point appeared subsequent to the 19[???] disallowance of personal interest. See [cite Taxpayer Advocate Reports].

The first component of the OID model (accrual at present value), posing an issue of timing, implicates the realization principle, but this issue is not critical (except where valuation is difficult), because a current deduction in the amount of the present value of a future expense is, standing alone, the equivalent of cash-method treatment. What is critical to the OID model is the combination of accrual at present value and the later deduction of the implicit interest. Combination of these two features results in the “double deduction” of interest, or “negative investment” treatment. In Part III it was shown that “negative investment” treatment of debt-financed expenses is economically unsound. It follows that obligations to pay future deductible expenses should be deducted when actually paid. This Part adopts a “realization” analysis to reach the same conclusion.

A. Mark-to-Market Treatment of Liabilities

The argument for the double deduction universal application of the current paradigm is ultimately based on the Simons income concept. It is commonly understood that the influential concept of income developed by Henry Simons (building on antecedent contributions by Schanz and Haig) was, in its ideal form, a mark-to-market (MTM) system under which the tax base reflected annual changes in net fair market values of assets. The Simons income concept is given normative underpinning by the notion of economic efficiency as specified by the Samuelson thesis, which holds that correct Simons (MTM) income measurement produces an outcome in which the effective tax rate is equal to the nominal tax rate. In the absence of correct income measurement, taxes would necessarily distort investment decisions by distorting investment values, as equal investments (before tax) would be taxed unequally.

The logic of Simons income measurement is that all future liabilities should generate double deductions, just as all future cash flow rights should be subject to double inclusions. To illustrate, assume the taxpayer incurs a fixed liability to pay $10,000 in five years, with no stated interest. Valuation theory posits that the future liability causes current wealth to be decreased by the current value of $10,000, which (on account of the time value of money) has to be significantly less than $10,000 (say, $7,473). Each year, the negative value of the liability will increase, on account of the passage of time, over the value as of the previous year, resulting in a negative value of $10,000 at the moment

---

168 Sunley, supra note ??? at 720, acknowledges the correctness of this conclusion in an off-handed way.  
169 An issue not discussed herein is whether the MTM (as opposed to realization) concept of income has a persuasive normative basis. It is not entirely clear what the normative basis of this concept was in the eyes of its creators, especially Simons (who appears to be hiding the ball in this respect). It may have been that objective market values were viewed as the best available proxy for subjective utility. Cite Haig ??? (income is “flow of satisfactions”). Intuitively, however, utility would seem more closely attached to the use of wealth rather than its acquisition or possession. Another possible (non-efficiency) normative basis for an MTM approach is that the MTM measure is the best base-line for redistribution. See Simons, supra note ???, at ??.
170 See note ??? [Samuelson]. As stated in this fashion, the Samuelson thesis is circular. It can also be claimed that there is no such thing as a “before tax rate of return,” because without taxes (and government) investments would not be possible. For the sake of discussion, these difficulties will be ignored here.  
171 See Calvin H. Johnson, supra note ??? (Soft Money), at 1039-42.
prior to payment. In order to reflect income accurately, the result should be a Year 1 deduction of $7,473 (the real principal) and deductions for Years 2 through 6 aggregating $2,527 (the real interest). Although it is true that aggregate deductions are $10,000, the deductible Year 1 amount of $7,473 is the current value of the obligation to pay $10,000 in the future, and the $10,000 includes the $2,527 that is deductible (again) as interest. This “double deduction” of interest to the payor is the mirror of the “double taxation” of interest income to the payee assuming that the payee followed the same accounting approach in reverse (namely, capitalization followed by interest income accrual).

In one sense, the Simons income norm proves too much, because it would render separate accounting for particular assets and liabilities irrelevant: the net income of firms (however defined) would be the net change in the value of the firm over the year, because the market valuation of the firm as a whole would take into account all assets and liabilities of whatever nature considered as an integrated activity. Similarly, the annual income or loss of individual equity-holders should, ideally, be measured by the change in value of their equity holdings, as opposed to a passing-through of internal tax accounting results. The valuation of firms (and, derivatively, of equity interests) assumes a stream of positive and negative estimated cash flows within the firm. Thus, the acquisition of particular future payment rights and obligations over time would be expected to have no necessary effect on the value of the firm unless such amounts fell outside of the predictions used by the market to establish existing values.\footnote{172}

The current income tax does not follow a “firm” or “equity interest” MTM approach. The next best approach would be one that would impose accurate MTM accounting for all assets and liabilities taken separately.\footnote{173} In a first-best world, any improvement in any tax accounting rule would improve overall economic efficiency. However, the notion that the income tax embodies anything resembling a sound accounting system is a pipedream. Congress and the Treasury have knowingly prescribed numerous rules that deviate from correct accounting, mostly in a way that favors taxpayers by deferring income\footnote{174} and accelerating deductions.\footnote{175} In such a second-best world, moving only liability accounting rules closer to the MTM ideal is likely to decrease over-all economic efficiency.\footnote{176}

**B. Accommodating Liabilities to the Realization Principal**

\footnote{172} This point is given “math” content in the text at note ???.

\footnote{173} See William A. Klein, Tax Accounting for Future Obligations: Basic Principles, 36 Tax Notes 624 (Aug. 3, 1987) (laying out ideal accounting treatment of the transaction described in Mooney Aircraft Co. v. United States, 420 U.S. 400 (5th Cir. 1969); ??? (future nuclear plant decommission costs viewed as being fully capitalized into the cost of a nuclear plant, resulting in treating the accounting as being for a single asset with a fixed internal rate of return of the same order of predictability as a debt instrument).

\footnote{174} Not only are investment gains subject to indefinite deferral under the realization principal, but also (and a point that is often overlooked) the cash method of accounting operates to defer vast amounts of income from services and the professions.

\footnote{175} Cite MACRS expensing rules; Treas. Reg. § 1.263-??? (known in tax circles as the anti-\textit{Indopco} rules relating to costs relating to certain intangibles).

\footnote{176} Theory of the Second Best ???
Universal MTM accounting at any level is considered to be non-viable for several reasons. First, it would violate the common understandings of ordinary people as to what gains and losses are “real.” Second, fairness and efficiency concerns would be posed by the taxation of speculative and non-liquid items. Third, the problem exists of ascertaining what is a separate asset or liability that should be treated as a distinct MTM item. Fourth, individual taxpayers cannot be viewed as MTM items, because individuals (as a “whole economic person”) are not bought and sold in the market; moreover, the tax system systematically ignores what is typically the person’s most valuable economic “component,” human capital, which is accounted for neither by references to changes in net value nor under more traditional accounting approaches.

The problems that would are posed by the foregoing are collectively dealt with under the current income tax by the realization principle, which is not a single rule but rather a collection of rules tailored to differing situations. Thus, changes in values of assets are generally not accounted for until sale or non-gratuitous disposition. However, some assets and liabilities are accounted for in advance of sale or disposition under OID, depreciation, and accrual rules. The issue can now be framed as follows: under a “realization” income tax, what is the appropriate tax treatment of future payment obligations that do not bear market-rate interest (the latter being treated as “borrowings”)? To attempt to answer this question requires a discussion of tax accounting methods as competing “realization” systems.

1. Scope and use of accounting methods

In the context of the present discussion of future payment obligations (not resulting from borrowings), the three relevant currently-available doctrinal approaches implementing the realization principle are the OID method, the cash method, and the accrual method. The OID rules, where applicable, supersede the others. Outside of the OID domain – which is quite narrow – the accrual method is mandatory for certain (business) taxpayers and activities, but not for professions and unincorporated services business. Where the accrual method is not required, the taxpayer elects (in effect) to follow the accrual method or the cash method on a consistent basis. In practice, the cash method is used by (a) individuals for their non-business activities, (b) professions (regardless of business form), (3) closely-held agriculture business, and (4) services businesses (not following GAAP accounting principles).

2. The cash method

177 Salience ??
179 For a recent discussion of the possibility of accretion accounting for changes in human capital values, see Kirk Stark, ??
180 Thus, human capital acquisition costs are not capitalized, resulting in an inability to implement a capital recovery (depreciation) system. See Joseph M. Dodge, ??.
181 See I.R.C. §§ 447, 448; Treas. Reg. § 1.446 ??.
182 Assuming no superseding mandates, if the taxpayer keeps its books according to GAAP accounting principles, it is on the accrual method; otherwise, it is on the cash method.
Under the cash method, liabilities of all kinds are deducted (or capitalized) only when paid (and in the amount paid). Deduction (or capitalization) does not occur when the liability is incurred or becomes foreseeable. The payment amount is deemed to be wholly “principal,” with no implicit interest component.\textsuperscript{183}

A proxy for “straight” cash method accounting would be a deduction, at the time the liability is incurred, of the present value of the future payment, but with no further deduction on account of actual payment.\textsuperscript{184}

The cash method is obviously far removed from a Simons income tax, because it ignores rights and obligations to future cash until they are actually paid.\textsuperscript{185} Indeed, the cash method sometimes ends up yielding the same result as under a consumption tax.\textsuperscript{186}

3. The OID method

The structure of an OID rule is similar to an MTM rule. Take, for instance, the example noted above of a liability to pay $10,000 in five years having a current negative value of $7,473. By adopting the convention that the discount rate one would use to obtain $7,473 as the present value of $10,000 – which happens to be 6% compounded annually - is to be applied on a constant basis to calculate the (compound) interest accruals to maturity, one has hit precisely upon the OID method of accounting.\textsuperscript{187} The OID rules in the Code specify a discount rate keyed to the most conservative investments. In so operating, possible market discounts relating to credit risks and imperfect markets are ignored.\textsuperscript{188} Despite the use of a constant interest rate and an excessively low discount

\textsuperscript{183} Under I.R.C. § 483, a cash payment for property might be considered to be implicit interest in part. However, transactions within § 483 typically involve capital expenditures and not expenses.

\textsuperscript{184} See H.R. Rep. No. 432, pt. 2, 98th Cong., 2d Sess. 1254 (1984) (this point is made in explaining the enactment of current § 461(h)).

\textsuperscript{185} On the gross income side, the cash method potentially conflicts with the principle that the receipt of property (in kind) is gross income at the FMV of the property on receipt. See, e.g., I.R.C. §§ 74 (prizes and awards); 83 (property received by service providers as compensation); Treas. Reg. § 1.61-1(a) (generally). A future payment right is considered property for most purposes, but (generally) not under the cash method. See Treas. Regs. § 1.83-3(e) (rights to future cash generally not “property” under I.R.C. § 83), 1.451-1(a) (cash method generally). The cash method contains exceptions for “cash equivalents” (checks, certain notes, and credit card transactions), but giving one’s own note is not a cash equivalent on the deduction side. See note ??.

\textsuperscript{186} Suppose an employer offers an employee a choice between $1,000 now and $1,123.60 ($1,000 plus 6% interest compounded annually) after two years. Under deferral, the interest will only be taxed once. In contrast, if the $1,000 were included by the employee in year 1 and then invested, it would be taxed twice. The employee is effectively obtaining consumption tax treatment for investment income. This result may be relatively harmless if the employer is also on the cash method and thereby simultaneously foregoing a double deduction of interest. See I.R.C. § 404(a)(5). The problems that would arise if the payor could accrue a $1,000 deduction in Year 1 are discussed in Daniel Halperin, The Time Value of Money, 23 Tax Notes 751, 753-55 (May 14, 1984).

\textsuperscript{187} Thus, the interest accrued in the first 12-month period would be 6% of $7,473 ($448), the (compound) interest for the second period would be 6% of the sum of $7,473 and $448, and so on.

\textsuperscript{188} See I.R.C. § 1272(a)(3) (original issue discount is internal rate of return compounded semi-annually). Credit risks in arms-length transactions are likely to be factored into (thereby increasing) the interest rate.
rate, the OID approach can be said to be the closest available proxy to MTM accounting under a realization system. The OID approach complies with the realization principle by reckoning “partial” gains and losses that accrue solely with the passage of time and are, therefore, “permanent and irreversible,” which is a necessary condition for any rule that deems realization to occur before sale or disposition.\(^{189}\)

The existing Code OID rules have very limited application, however. Basically, they are imposed only with respect to certain obligations to pay a fixed-dollar cash amount at a specified future date\(^{190}\) where the stated interest rate is zero or “below market”\(^{191}\) and where (on the income side) the context is deemed to pose an opportunity for tax avoidance.\(^{192}\)

It may be thought that the narrow scope of the OID rules is attributable to “math phobia.” Perhaps, however, this reason is over-rated, as business taxpayers and their hired guns would be able to handle the calculations, either through published tables (similar to the actuarial tables already used in the estate and gift tax) or by way of computer spread-sheet programs (and even hand-held calculators). The more important reason for the narrow scope of the OID approach is that only certain financial instruments (providing for fixed dollar amounts at specified future dates), if accounted for under the OID approach, produce reliable numbers that do not have to be corrected retroactively. Thus, in addition to the “permanent and irreversible” criterion of realization (before disposition), there appears to a separate “accuracy” requirement.\(^{193}\)

Of course it can be claimed that OID inheres in all assets and liabilities (anything in which future rights and obligations can be reduced to present value), or, to state it in more technical terms, that all assets and liabilities have an internal rate of return that can be measured.\(^{194}\) Perhaps such is true in hindsight; moreover, the market (and firm

---

\(^{189}\) See I.R.C. §§ 165(a) (loss must be “sustained” if not resulting from sale or disposition) and 167(a) (depreciation deductions only for assets that are used up with the passage of time); Treas. Reg. § 1.165-1(b) and (d)(2) (sustained loss requires closed and completed transaction, fixed by identifiable events, and without a reasonable prospect of recovery).

\(^{190}\) OID is the excess of the “stated redemption price at maturity” over the “issue price.” See I.R.C. § 1273(a) & (c). Passage-of-time gains that occur with respect to interests in property (remainders, reversions) are ignored. The OID rules presuppose that the consideration received by the obligor be fixed, as is the case with cash loans and obligations to pay for property (with an ascertainable value); obligations to pay for services are exempt from the OID rules unless the obligations are publicly traded. See I.R.C. §§ 1273(b) and 1274.

\(^{191}\) However, a parallel set of rules applies for bonds providing for above-market interest rates. See I.R.C. § 171; Treas. Reg. § 1.163-?? (bond premium).

\(^{192}\) Thus, I.R.C. § 1272 counts inconsistent treatment of issuer and holder of an OID obligation; I.R.C. §§ 483 and 1274 counter an opportunity to convert interest income into capital gains in deferred-payment sales; and, I.R.C. § 7872 counts disguised economic transfers (gifts, salary, dividends) between certain related or quasi-related parties. I.R.C. § 163(c) has long allowed accrual-method debt issuers to deduct OID as it accrues.

\(^{193}\) See Treas, Reg. § 1.446-1(c)(1)(ii) (prerequisite to accrual is that the amount must be determinable with reasonable accuracy).

\(^{194}\) See Johnson, ???.

---
managers) can be assumed to predict future events and discount them back to the present. However, these are arguments for an MTM system. The realization principal requires something more than estimates, predictions, and the ability to make calculations. In virtually all of its manifestations, it requires definiteness and finality. Any OID “buried” in an asset or liability can be swamped by changes in “outside” (non-passage-of-time) valuation factors. Disaggregating passage-of-time gains and losses from gains and losses attributable to changes in estimated future yields and discount rates is a nice analytic trick, but it is divorced from any kind of reality except in the case of fixed-payment financial instruments. To state the matter bluntly, the choice of a realization income tax over a Simons MTM income tax is a decision that implies confinement of the OID approach to cases involving an obligation to pay sums certain at specified dates. OID treatment of garden-variety trade accounts receivable and accounts payable, bearing no stated interest and having no fixed payment date, would be beyond the current OID realm. Even if they were closer to the margin, it would not be worth the effort to account for them on an OID basis.

Those who advocate OID treatment for liabilities (such as for nuclear decommissioning costs and environmental clean-up costs) ignore the far worthier candidates for OID treatment on the gross income side, namely, obligations acquired at a market discount, annuities, life insurance, remainder interests, reversionary interests, retirement benefits, and other income deferred through trust or escrow devices. All of these contain passage-of-time gains and some of them involve fixed-dollar monetary claims.195

Not only are the OID rules in the Code under-inclusive on the income side, but the principle that underlies the OID approach (realization solely due to the passage of time) is extended outside of the OID area in a way that serves solely to favor taxpayers on the deduction side. The reference here is to allowance of deductions for depreciation, amortization, and depletion for all business and investment assets, both tangible and intangible, except for those few asset types having an indefinite duration. The one-sided application of the passage-of-time concept is easy to explain. Appreciation has an unlimited upside potential, so that there is no mathematical formula that can treat market appreciation (or internal growth) in a way that resembles the OID approach of discounting a future amount back to the present. In contrast, (depreciable) assets always have a terminal value of zero. If not, as under the current Code, they can be viewed as having a terminal value of zero. The combination of a known cost, a terminal value of zero, and an estimated useful life based on statistics allows for the easy creation of an OID-like formula for computing passage-of-time losses that could have plausible accuracy if depreciation formulas were constructed correctly. However, they are not. Actual depreciation rules under the Code are more favorable to taxpayers than an OID formula in (1) simply assuming that all (tangible) assets have a terminal value of zero, (2) assuming shorter useful lives than industry experience, and (3) deviating from a constant-rate (OID) approach.196

195 At the same time, life insurance companies can accrue reserves for mortality risks. See I.R.C. § 807(d).
196 See Calvin H. Johnson, supra note ??? (Soft Money), at 1041-53.
Table 3 below illustrates the error in depreciation caused solely by abandonment of the constant-rate (OID) approach. The asset, costing $249, is a right to receive $100 annually for three years. The $249 cost is the sum of the present values ($91, $83, and $75) of three annual $100 payments using a 10% constant discount rate compounded annually. The economic depreciation is the decrease in the present value of the investment (remaining income stream) attributable to the passage of time. Thus, at the moment the year 1 payment is received, the remaining two payments have a present value of $174 ($91 + $83). The economic depreciation for Year 1, then, is $75 ($249 - $174). Economic depreciation is contrasted with straight-line (S-L) amortization, which produces a Year 1 deduction of $83.\(^{197}\) The MACRS column describes the depreciation under Code section 168 that would be available if the asset were “3 year” tangible personal property.”

### Table 3 – Economic vs. Tax Depreciation

<table>
<thead>
<tr>
<th>Time</th>
<th>PV of $100 rec’d in</th>
<th>Depreciation Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yr 1</td>
<td>Yr 2</td>
</tr>
<tr>
<td>Begin Year 1</td>
<td>$91</td>
<td>$83</td>
</tr>
<tr>
<td>End Year 1</td>
<td>-</td>
<td>$91</td>
</tr>
<tr>
<td>End Year 2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>End Year 3</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Economic depreciation can be described as (1) a loss from the severance from the investment of the Year 1 payment assigned a present value of $91, reduced by (2) the passage-of-time gain resulting from the advancement to maturity of the payments for years 2 and 3 ($25). The Year 2 payment increases in present value from $83 to $91 (= $8), and the Year 3 payment increases in present value from $75 to $83 (= $8) (total gain = $16) (net loss = -$91 + $16). The depreciation systems provided by the Code err in ignoring these passage-of-time gains.

The error illustrated by Table 3 is compounded by the failure to take terminal value (estimated salvage) into account. It is compounded further by assuming short useful lives. In sum, the depreciation system flunks the “accuracy” test for realization and is slipshod on the “finality” aspect as well, because depreciation is commonly taken on assets that actually appreciate in value or last much longer than the write-off period.\(^{198}\) It is not simply inconsistent to follow a lax standard of realization on the deduction side and the highest standard on the income side, it fatally undermines the integrity of the realization income tax as a whole.

\(^{197}\) S-L depreciation is calculated by dividing the cost by the useful life (3 years).

\(^{198}\) Examples are buildings and collectibles used in a business. Cite Liddle case ???.

198 Examples are buildings and collectibles used in a business. Cite Liddle case ???
It would be possible to apply a uniformly low standard of realization across the board, in other words, to allow formulaic depreciation deductions together with formulaic income inclusions. Indeed, the two operations can be combined by specifying an imputed net return rate. The net (after depreciation) return on an asset can be imputed simply by applying the same discount rate used to value the asset against the basis of the asset (as increased by prior year inclusions and reduced by receipts). This is shown in Table 4, using the same facts and 10% imputation (discount) rate as in Table 3.

### TABLE 4 – COMPARISON OF INCOME-IMPUTATION AND DEPRECIATION APPROACHES TO INCOME MEASUREMENT

<table>
<thead>
<tr>
<th>Basis</th>
<th>10% Return</th>
<th>Imputation Method Basis Adjustment</th>
<th>Economic Depreciation Method</th>
<th>Net Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>$249</td>
<td>$25</td>
<td>($+25 - $100)</td>
<td>$100</td>
</tr>
<tr>
<td>Year 2</td>
<td>$174</td>
<td>$17</td>
<td>($+17 - $100)</td>
<td>$100</td>
</tr>
<tr>
<td>Year 3</td>
<td>$91</td>
<td>$9</td>
<td>($+9 -$100)</td>
<td>$100</td>
</tr>
</tbody>
</table>

It can be seen that the imputation system described on the left side of Table 4 produces the same results as under economic (present value) depreciation. This is no coincidence. In fact both approaches are the same as what produces OID calculations. The calculation of economic depreciation is really the same as the calculation of constant-rate OID. Under the latter, the discount rate (10%), applied against the investment’s basis (original basis plus income inclusions less cash distributions), produces (net) income for Years 1, 2, and 3 of $25, $17, and $9 respectively. 199 This is the same set of results as are reached by subtracting the economic depreciation deductions from the $100 receipts. The same set of results is also obtained by viewing each future receipt as an OID obligation. 200

---

199 The constant rate (10%) is applied against the initial basis decreased by cash withdrawals and increased by prior income inclusions. Year 1: 0.1 x $249 = $24.9 (rounded up to $25). Year 2: 0.1 x ($249 - $100 + $25) = $17.49 (rounded down to $17). Year 3: 0.1 x ($249 - $200 + $25 + $17) = $9.1 (rounded down to $9).

200 The following treats each right to receive $100 in the future as a separate OID obligation:
The imputation method is the simplest as far as the math is concerned. It is also easiest to implement because it does not really depend on an accurate fix on the amount and date of future payments. Just as the OID rules use the AFR (“applicable federal rate”), which is the most conservative rate of return (based on the yield on U.S. government obligations), so can the imputation method use the AFR. Finally, the imputation method can be applied to depreciating assets, assets with no ascertainable useful life, and assets that necessarily appreciate with the passage of time.

There are various problems with a universal imputation system, despite its internal coherence. One is that implementation of the imputation method depends on the ability to identify assets and liabilities and to assign them an initial (positive or negative) basis equal to initial (positive and negative) value. Thus, the imputation method is dependent on a pervasive system of capitalization of costs (and of reverse capitalization of future costs). Unfortunately, the capitalization principle on the asset side is currently more honored in the breach than in the observance. In addition, identifying the cost of the many important intangibles is extremely difficult. On the deduction side, liabilities (other than borrowings) that would be recognized under an imputation system are not recognized at all under the cash method, and many are also not recognized under the accrual method, discussed below. Moreover, an item-by-item approach to imputation is probably unworkable on account of the fact that it would be impossible to allocate income receipts among the various assets comprising a business.

The problem of having to accurate account for each asset and liability could be finessed by fusing the whole bundle into a single asset (the firm or activity). However, an MTM system would be much simpler and, by definition, more accurate, except for cases where valuation of a firm is difficult. In the academic literature, imputation accounting has only been proposed for discrete investments that neither generate conventional current income nor can be feasibly marked to market. However, imputation accounting for high-risk investments, like unproven intangibles or mineral interests, would be strongly resisted on the grounds that treating such investments like government bonds would be highly inappropriate. For the same reason, the government would resist imputation accounting for contingent liabilities.

The argument above might be characterized as a variation of that made in the context of an MTM system. OID or imputation treatment of liabilities would be incorrect unless all (or at least most) assets and liabilities were taxed on an imputation basis. Stated differently, imputation with respect to assets and liabilities is necessary in order to apply the lax realization standard of depreciation across the board. The alternative is to conform depreciation rules to OID standards. However, neither reform is likely.

4. The accrual method

---

201 The imputation system would operate like the accrual method. Referring to the table in the preceding footnote, in Year 1 there would be total income inclusions of $25, and the basis of each payment right would be increased accordingly. The receipt of the Year 1 payment would not produce additional income, because the basis would be $100 (initial basis of $91 plus income inclusion of $9). The basis of the rights to the payments for Years 2 and 3 would also have a basis of $100 upon receipt.

202 See Mary Louise Fellows. ????
The third doctrinal approach to realization in the business context is the accrual method, which is actually a bundle of rules that allow only for a single deduction for the face amount (rather than the present value) of the obligation but which often allows the deduction to be taken in advance of actual payment. On the liability side, the accrual method allows the current deduction of the face (stated principal) amount of a future payment obligation upon the satisfaction of both the all-events test and the economic-performance test. These tests operate to delay accrual. The all-events test requires that all the events have occurred that trigger the liability, provided that the amount of the liability be reasonably ascertainable. Over time, the all-events test has been tightened so as to preclude accrual for contingent (and disputed) liabilities as well as for liabilities that can be predicted only on the basis of statistical experience. All of these moves have distanced the accrual method from a Simons income (MTM) tax, because future liabilities (no matter how contingent and speculative) would affect the valuation of the firm. However, such liabilities flunk the realization standards of “finality” and or “accuracy.”

The economic-performance test further delays accrual (at face amount) until whichever is applicable of: (1) the year (prior to payment) that the obligation is owed (earned) by reason of taxpayer use, (2) the year (prior to payment) that the obligation is earned by vendor performance, (3) the year (prior to payment) that the obligation is satisfied by taxpayer performance, or (4) the year that the obligation is actually satisfied by cash payment.

In my view, the economic-performance test was a long-overdue correction of the all-events test. Previously, the entering into of a binding contract to pay (or expend) a fixed amount upon future performance by either party had satisfied the all-events test on the theory that the liability was “fixed” in the sense of being certain and predictable. However, in the economic sense the liability is not sufficiently fixed until the obligation to pay is earned by performance. Thus, it had long been understood that rent and interest

---

203 On the income side, accrual occurs upon satisfaction of only the all-events test.
206 Thus, in United States v, General Dynamics, 481 U.S. 239 (1987), the Court held that a taxpayer could not accrue deductions for self-insured workforce health care costs until claims were filed (as opposed to the earlier time when health costs were incurred), on the ground that the filing of the claims was by no means automatic.
207 See I.R.C. § 461(h)(2)(A)(iii). This rule covers interest, rent, and sometimes property taxes.
208 See I.R.C. § 461(h)(2)(A)(i) & (ii) (services and property provided to the taxpayer). This rule covers most accounts payable, but not deferred compensation, which is deducted when paid pursuant to other Code provisions.
209 See I.R.C. § 461(h)(2)(B). This rule covers warranty obligations.
210 See I.R.C. § 461(h)(2)(C) & (D); Treas. Reg. § 1.461-4(g) (most taxes, tort and worker compensation claims, prizes, jackpots, rebates, refunds, permit fees).
211 See, e.g., Ohio River Colleries v. Comm’r, 77 T.C. 1369 (1981) (accrual of future mining reclamation expenditures). Accrual was not always based solely on legalities; accrual was sometimes held to occur upon the existence of facts that underpin a future liability (such as for income taxes). See Anderson, discussed in the text at note ??
accrues only as it is earned by the provider (with the passage of time), not when loan agreements and leases are entered into. The economic-performance test operates to conform the rest of accrual doctrine to the “earned” principle governing interest and rent.

The reason why “earned” should be a prerequisite for accrual is best understood by considering business accounts payable. Until performance by the vendor, there has been no transfer of value to the taxpayer, and the vendor would have no reason to charge interest (or rent). If there is no basis for an interest charge prior to performance, there is no real basis for treating the “liability” as if it were a borrowing or a debt-financed expenditure. And, if there is no borrowing, there is no expenditure. As has already been shown, the accrual of a cost in advance of its expenditure creates a local tax shelter effect.

C. Should the Accrual Method Be Abolished?

Herein it will be assumed that a liability should not be accrued any earlier than the satisfaction of all of the following: (1) the liability is “fixed,” (2) it is “earned,” and (3) the amount can be determined with sufficient accuracy. Even then, accrual at face (in advance of its payment) can only be justified if the liability bears market-rate interest. In that case, the accrual of the deduction at face can be viewed as the borrowing of a principal amount equal to the face amount coupled with an expenditure of such amount. If market-rate interest is not charged, then accrual at face gives an even more favorable result to taxpayers than under the OID approach. This point should be intuitively obvious: a deduction of $10K in advance of payment is worth more (assuming a constant tax rate) than a deduction of, say, $8K ($10K reduced to present value) now and $2K later (as under a strict OID approach). Basically, accrual at face here overstates the amount expended as an expense or capital expenditure. Accrual at face is presumably tolerated only because it also occurs on the income side. But two wrongs do not make a right unless there is no satisfactory alternative. One alternative is the OID approach, but that presupposes a fixed payment date as well as a fixed amount. Moreover, applying the OID method is probably not worth the effort for routine trade accounts payable and receivable. The question, then, is whether the cash method is preferable to the accrual method.

This latter question may not be easy to answer, because, if the OID approach is theoretically correct, then the choice between the accrual and cash methods is a toss-up, neither being correct, and the tie-breaker might be the convenience of substantial conformity to book accounting or the ease of cash accounting. The academic literature treats the implicit-interest (OID) scenario as being obviously correct in principle. This claim will now be addressed on the merits.

---

212 See text at note ???.
213 See Kiefer, note ???, at 927 (deferred payment example assumes that interest is charged, either implicitly or explicitly).
The OID approach converts a non-market-interest-bearing right or obligation into an obligation that bears implicit interest on a “real” principal that is less than the stated principal (face amount), because the real principal is the present value of payments under the contract. The claim that all future payment obligations (and rights) contain an implicit interest component strikes me as being dogmatic, because it is really a claim that all delays in payment necessarily entail implicit interest solely because the future payment can be reduced (by arithmetic) to present value. A similar claim would be made on the income side for future payment rights. The math is easy if the future payment date and amount are fixed, but even if they are not, the implicit interest can be calculated retroactively once the payment date and amount are known by virtue of hindsight.\(^{214}\)

However, just because a present value can be calculated does not lead to the conclusion that there is any interest, defined as compensation for the use of money.\(^{215}\) To delineate the issue here as narrowly as possible, three classes of cases (all mentioned earlier) can be immediately identified as being devoid of implicit interest: (1) cases involving market-rate stated interest, (2) cases involving the purchase of goods or services that have not yet been received by the taxpayer, and (3) cases where the taxpayer’s obligation relates to something other than the payment of money.\(^{216}\)

In cases where goods and services have been received by the taxpayer, and the liability is not contingent (etc.), it can be said that there is a realization event sufficient to cause accrual in an amount equal to the present value of the liability. But the item accrued can be either an expense or a capital expenditure. It was shown in Part III that “borrowing” treatment – which an OID approach would achieve – is appropriate only if the (real) principal is invested in the tax sense, i.e., treated as a capital expenditure, because only in that case is the “double taxation” of income matched by the “double deduction” of interest. Indeed, the Code already provides for implicit interest in the case of most deferred-payment purchases of property.\(^{217}\) On the other hand, OID treatment of an item deducted as an expense creates a tax shelter that is unjustified by the economics.\(^{218}\) Thus, interest should not be imputed in such a case, and the expense deduction should either be taken currently in an amount equal to the present value of the future cost, or, if the present-value computation is subject to infirmity, the deduction

---

\(^{214}\) The retroactive computation of implicit interest is required for most contingent-payment sales of property. See ???.

\(^{215}\) Commentators typically assume the existence of an interest charge by stipulating a deferred payment amount in excess of the value of the goods or services. See Halperin, supra note ??? (Interest), at 516-19; Kiefer, note ???, at 926-27. However, value does not exist in the abstract independently of the market price, and the fact of deferred payment does not per se lead to the conclusion that the market price at the time the goods or services are provided must necessarily be an amount equal to the present discounted value of the amount paid. The assertion to the contrary is dogmatic because it cannot be empirically disproven.

\(^{216}\) See Calvin H. Johnson, note ??? (Earned), at ??? (arguing convincingly that prepaid income for services or goods to be provided by the taxpayer in the future is not an interest-bearing loan or the equivalent thereof). In cases of this type, the taxpayer has discretion over the means (and the cost) of providing the goods and services.

\(^{217}\) See I.R.C. §§ 483, 1274.

\(^{218}\) See text at note ???.
should be taken only when it is paid. These are alternate specifications of the cash method.

Certain items (like inventory purchases) lie on the borderline between expenses and capital expenditures. Here it would seem that implicit interest might be said to exist if the price increases according to the length of the delay in payment. For example, the well-known *Mooney Aircraft* case\(^{219}\) could not have involved implicit interest, since the payment obligation did not increase with the passage of (a long period of) time and the payment date was under the payee’s control. Other factors that might be indicative of the absence of imputed interest\(^{220}\) include (1) the absence of a specified future due date(s) for payment (relative to the date of performance), and (2) the unsuitability as collateral of the item purchased (examples would include, fuel, agricultural produce, and supplies). Most trade payables are due upon the rendering of a bill, and it is typical for no interest charges to be specified in the billing statement (or, if they are stated, for interest not to commence running for some meaningful grace period).

It might be said that vendors in commerce expect that some customers will not make timely payment, and that therefore an interest charge (as well as a charge for the risk of nonpayment) is built into the price of goods and services. The flaw in this argument is that, in the absence of stated interest, this charge cannot be for any actual delay in payment of a particular vendee. Suppose customers A and B are charged the same price, but A pays on time, whereas B pays one year late. It cannot be said that part of what A pays is “interest,” because A has not delayed payment. Since B pays the same price as A, it is equally hard to say that B is paying interest. Both of customers A and B are just paying the market price, which is as high as the traffic will bear. There is no reason to think that the vendor is charging any more to its customers than it would charge if all were to pay in a timely fashion. The fact that the vendor does not charge interest amounts to a subtle kind of price discrimination in favor of B, which (as an effective price decrease) is the negation of interest. Even discounts for cash purchases might be viewed as a form of price discrimination or a discount for avoiding transaction costs rather than as implying that interest exists on non-cash sales.

Thus, doubts about the existence of interest combined with the tax-shelter effect of OID treatment of debt-financed expenses (and inventory costs) argues for cash-method treatment. The same logic applies on the income (receivables) side.

That leaves for consideration the case where the future liability arises from neither the receipt nor the provision of goods and services. Here there is no investment by the taxpayer that commands either current accrual of an expense – there having been no “cost” – or the imputation of interest expense. A case where there was clearly no investment by payor (or payee) was the *Hughes Properties* case.\(^{221}\) There the future

---

\(^{219}\) *Mooney Aircraft Co. v. U.S.*, 420 F.2d 400 (5th Cir. 1969) (fixed dollar amount payable on retirement of sold aircraft).

\(^{220}\) These factors are adapted (in part) from the tax doctrine pertaining to the distinction between debt and equity. Cite Boris Bittker and James Eustice ???.

payoff on a progressive slot machine grew with the passage of time, but the payoff was not deferred with respect to any particular customer with respect to any particular prior investment of the payee.

However, there are some liabilities that can be characterized as containing an implicit interest element. A case of this sort was the Ford Motor Company case,\textsuperscript{222} involving a structured tort settlement. However, the existence of implicit interest does not justify OID treatment, because the latter is really the same as borrowing treatment, and a deemed borrowing assumes some use (spending) of the funds deemed to have borrowed. In a case like Ford Motor Company the taxpayer obtained nothing of identifiable value in return for incurring the tort liability. Thus, the theoretically correct tax treatment in that type of case would be to treat the implicit interest as interest but to delay the deduction of the real principal until payment. However, it is probably not worth the effort to distinguish the interest expense deduction from the principal expense deduction. The entire amount can be deducted on the cash method.

An additional reason for cash-method treatment in this kind of case comes into play where the payee is on the cash method (and defers inclusion). Here accrual in any form would create a tax accounting mismatch that could be exploited by the parties at the expense of the Treasury.\textsuperscript{223} The Code already disallows accrual in many instances of this sort.\textsuperscript{224} However, the dynamic does not operate in reverse. That is, even if the payee were “double taxed” as an investor, that would not justify treating the payor as a negative investor.\textsuperscript{225}

In sum, it turns out that the OID approach is not the norm for most cases that are presently considered to be in the “accounting methods” domain, which deals with the timing of expense and gross income items. Credit-financed capital expenditures are distinguishable, of course. It appears, then, that the accrual method “as we know it” should be abolished!\textsuperscript{226} Indeed, it has been abolished in the case of most taxes, fines, tort liabilities, jackpots, refunds, and the like. The cash method should be extended to all cases not clearly involving “investment.”

D. A Case Study: Nuclear De-Commissioning Costs

A good deal of the discussion took place in the mid-1980s concerning not only the tax treatment of accruals generally\textsuperscript{227} but specifically the tax treatment of nuclear

\textsuperscript{222} Ford Motor Company v. Comm’r, 71 F.3d 209 (6th Cir. 1995).
\textsuperscript{223} See Calvin H. Johnson, Silk Purses ???
\textsuperscript{224} See I.R.C. §§ 267(a), 404(e), and 467. ???
\textsuperscript{225} An analogy would be paying wages for house cleaning. That the wages are includible by the payee does not justify a deduction to the payor.
\textsuperscript{226} Cash-method treatment would also apply to inventory costs and capital expenditures in cases where there is no stated or implicit interest.
\textsuperscript{227} An entry at 26 Tax Notes 941 (March 4, 1985) lists 17 articles, reports, and comments published in Tax Notes from late 1982 through early 1985.
decommissioning costs. A firm undertaking the building of a nuclear power facility undertakes an expensive obligation to leave a decontaminated site when the facility ceases to operate. This termination cost exceeds any conventional positive salvage value, producing net negative salvage. A discussion of this issue summarizes many of the points made earlier. For discussion purposes, it will be assumed that these costs can be estimated with reasonable accuracy and economic depreciation (and appreciation) is the computational norm.

The problem is posed herein by considering a negative-salvage hypothetical that is embedded in an overall economic wash that should, if accounted for correctly, produce a tax wash. The hypothetical is an investment that will produce a positive cash flow of $100 for three years and a termination cost after three years of $331. At a 10% discount rate, the present value of both the positive and negative cash flows is $249. Accordingly, the investment is worth zero ex ante. It is assumed that the $100 Year 1 and Year 2 receipts are reinvested at 10% in order to create the possibility that the investor will end up “even” ex post. Table 5 below outlines the economics of the proffered hypothetical, which is really an offsetting asset-liability pair, in present value terms.

TABLE 5 – OFFSETTING ASSET AND LIABILITY (BEFORE TAXES)

<table>
<thead>
<tr>
<th>Asset Side</th>
<th>Liability Side</th>
<th>Net Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV Rec’t</td>
<td>Exp.</td>
<td>Gain</td>
</tr>
<tr>
<td>PV Rec’t</td>
<td>PV Exp.</td>
<td>OID</td>
</tr>
<tr>
<td>PV Rec’t</td>
<td>Gain</td>
<td>Net PV Loss</td>
</tr>
</tbody>
</table>

The depreciation is economic depreciation. The “Gain” column refers to the annual before-tax profit (total cash income from the investment and the reinvested cash less economic depreciation and less implicit interest expense). The Net Value of the investment (PV Rec’t less PV Exp.) starts at zero, due to the offsetting present values, and decreases over time because the lump-sum expense occurs at the end, but the decrease is offset by the reinvested receipts augmented by compound interest. At the end, and assuming there is no consumption, the taxpayer before taxes has no overall gain or loss. There is an aggregate investment PV loss of $331 that is accounted for by depreciation deductions of $249 and interest deductions of $82, and aggregate cash GI is also $331.

Before considering this hypo as an investment in a nuclear facility, consider the more straightforward possibility that the taxpayer borrows $249 either from a third party or the seller to finance the asset purchase for $249. The loan, plus accrued interest, is to be paid off at the end of three years in a lump sum. Note that the liability is not built into the internal cash flow of the asset, so that the future cost would not be capitalized into the price of the investment.

---

228 See Sunley, note ???, ??
TABLE 6 – TAXATION OF DEBT-FINANCED ASSET ACQUISITION

<table>
<thead>
<tr>
<th></th>
<th>Receipt</th>
<th>Int.</th>
<th>PV Rec’t</th>
<th>Deprec.</th>
<th>Principal</th>
<th>OID.</th>
<th>TI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begin</td>
<td>0</td>
<td>0</td>
<td></td>
<td>249</td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>End Year 1</td>
<td>+100</td>
<td>0</td>
<td>174</td>
<td>-75</td>
<td>0</td>
<td>-25</td>
<td>0</td>
</tr>
<tr>
<td>End Year 2</td>
<td>+100</td>
<td>+10</td>
<td>91</td>
<td>-83</td>
<td>0</td>
<td>-27</td>
<td>0</td>
</tr>
<tr>
<td>End Year 3</td>
<td>+100</td>
<td>+21</td>
<td>0</td>
<td>-91</td>
<td>-249</td>
<td>-30</td>
<td>0</td>
</tr>
</tbody>
</table>

Since the loan provides for no stated interest payments along the way, it is an OID obligation, with implicit interest as indicated. The borrowed money is excluded and the capital expenditure is not deducted but provides either an initial cost basis or a *Crane* basis of $249. Each year, the sum of the economic depreciation deduction and the interest deduction exactly equal the gross income. The $249 payment at the end is a non-deductible principal repayment. Hence, there is no tax along the way, and the compounded positive cash flow of $331 is exactly equal to the terminal payment of $331. The terminal basis is zero on account of the depreciation deductions.

Table 6 confirms that proposition that the current tax treatment of debt is the correct match for correct income tax asset accounting.

Next the hypo will be dressed in the clothing of a nuclear power facility, and the crucial assumption is future cost is inherent in the investment itself, and not the result of a current borrowing. Thus, the payment at the end is an expense, not a loan principal repayment. Here the future cost is capitalized into the investment itself, which would have a beginning net value (and cost) of zero. (In effect, the vendor would be financing the future cost.) It follows from the fact of a zero purchase price that there are no depreciation deductions.

Table 7 below assumes IRA (consumption tax) treatment of the investment. That is, all gross income receipts are reinvested and deducted, or the economic return takes the form of unrealized appreciation, so that the final withdrawal ($331) is taxed with zero basis offset. The issue is what tax treatment on the cost side produces the best result, and the obvious answer is a year 3 “expense” deduction of $331.

TABLE 7 – INTERNALIZED COST (CONSUMPTION TAX MODE)

<table>
<thead>
<tr>
<th></th>
<th>Receipt</th>
<th>Int.</th>
<th>Income</th>
<th>Payment</th>
<th>Deduction</th>
<th>Net gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begin</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>End Year 1</td>
<td>+100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>End Year 2</td>
<td>+100</td>
<td>+10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>End Year 3</td>
<td>+100</td>
<td>+21</td>
<td>+331</td>
<td>-331</td>
<td>-331</td>
<td>0</td>
</tr>
</tbody>
</table>

Because of the consumption tax treatment of the positive cash flow, the correct deduction rule is cash-method treatment of the negative cash flow. There would be no OID accruals. This conclusion is a reiteration of the point made earlier that income tax
(OID) treatment of liabilities should not be matched with consumption tax treatment of assets.

In Table 8 below, the zero-net-cost scenario is subjected to income tax treatment, with the tax rate assumed to be 20%. To simplify the math, it is assumed that tax payments and refunds are out of (or into) a separate fund, rather than affecting the amount the investment itself. The issue here is how to account for the obligation to pay a $331 cost in the future. The alternatives are many: (1) deduct $331 initially (old-style accrual treatment); (2) deduct $249 initially and deduct the implicit interest expense as it accrues; (3) deduct the implicit interest expense as it accrues and deduct an expense item of $249 at the end; and (4) deduct $331 at the end (cash method treatment). Table 8 incorporates the third of these methods as the best available under the circumstances.

### TABLE 8 – INTERNALIZED COST (INCOME TAX MODE)

<table>
<thead>
<tr>
<th>Receipt</th>
<th>Int.</th>
<th>Expense.</th>
<th>OID</th>
<th>TI</th>
<th>Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begin</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>End Year 1</td>
<td>+100</td>
<td>0</td>
<td>0</td>
<td>-25</td>
<td>+75</td>
</tr>
<tr>
<td>End Year 2</td>
<td>+100</td>
<td>+10</td>
<td>0</td>
<td>-27</td>
<td>+83</td>
</tr>
<tr>
<td>End Year 3</td>
<td>+100</td>
<td>+21</td>
<td>-249</td>
<td>-30</td>
<td>-158</td>
</tr>
</tbody>
</table>

The TI for Years 1 and 2 is the income receipts less the OID expense accruals. The terminal payment of $331 is treated as a business expense of $249, an interest (OID) expense of $30, and a nondeductible payment of $52 (representing prior OID deductions). Total income is $331, as are total deductions.

Alternatives (1) and (2), both providing for up-front deduction accruals, are wrong, because such accruals would drastically overstate the Year 1 loss in the net present value of the investment. Alternative (4) also is wrong by treating the asset loss as occurring all in Year 3. That leaves alternative (3) as the closest approximation to the correct tax result of the four alternatives offered. The correct tax result is the same result as shown in Table 5 as the Net Present Value Loss and in Table 6 as the sum of Depreciation and Interest Deductions. However, the depreciation shown is economic depreciation, which doesn’t exist in the Code (except for debt instruments other than annuities).

Table 9 below shows alternative sets of taxable income (TI) results under (1) straight-line three-year amortization and (2) three-year MACRS (200% D-B) depreciation, which are the principal depreciation methods under the Code.\(^{229}\)

### TABLE 9 – DEPRECIATION AND OID

\(^{229}\)Intangible assets are amortized over their duration on a straight-line basis under I.R.C. § 167, and tangible assets are depreciated under the 200% declining-balance method (changing to straight line at about the mid-point), with a half-year convention.
To sum up, where the future liability is capitalized into the cost of the investment, the theoretically correct treatment is to treat the investment as a unitary OID obligation, so that each year the net income (or net loss) would equal the change in the net present value of the investment.\textsuperscript{231} Of course, where an investment of this type is accounted for correctly, it would be incorrect to provide for an additional set of deductions for the future liability, because that would result in a double deduction of the same cost.

However, the kinds of investments subject to OID treatment under the Code (certain debt obligations) are not of the type that have built-in costs. In general, the class of assets that might conceivably reflect capitalization of expected future costs (and revenues) would be equity interests in businesses, but these have no “stated redemption price at maturity,” and therefore cannot qualify for OID treatment under the current realization income tax. Even if the current norm of realization were relaxed to allow for OID treatment of equity interests, applying the OID (change in PV approach) to an equity interest would involve so much guesswork as to not be worth the effort, and the OID approach would surely yield to the first-best MTM approach.

If nuclear de-commissioning costs are not fully capitalized into the price of a nuclear plant, then a “unitary” accounting approach is not feasible. Full capitalization seems unlikely, because an electric utility would presumably construct its own plant. Therefore, and the price of the inputs would not be reduced to reflect the decommissioning cost, because these inputs would have non-encumbered uses to other buyers.\textsuperscript{232} In the absence of full capitalization, the asset and liability sides of the investment would have to be accounted for separately. The liability would not actually qualify for OID treatment, because the nuclear decommissioning costs are not a fixed dollar obligation payable at a specific date. Finally, the assets would be depreciated separately under an economically incorrect accelerated method.\textsuperscript{233} Therefore, OID treatment of the liability would be too generous.

Congress has not accepted the OID approach to decommissioning costs, but has adopted a variation of the cash method.\textsuperscript{234}

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
Receipt & Int. & OID & S-L Deprec. & S-L TI & 200% D-B & D-B TI \\
\hline
End Year 1 & +100 & 0 & -25 & -83 & -8 & -83 \hline
End Year 2 & +100 & +10 & -27 & -83 & 0 & -110.67 \hline
End year 3 & +100 & +21 & 30 & -83 & +8 & -55.33 \hline
\end{tabular}
\end{table}

\textsuperscript{230} The remaining basis is deemed to be taken as depreciation in the year of disposition.
\textsuperscript{231} Cite ??? In cases where the future cost is capitalized as a reduction in the cost of the investment, that future cost must be a necessary condition to obtaining the positive cash flow of the investment, so that such cost should be charged against the revenue received in advance of paying the cost. See text at note ???.
\textsuperscript{232} Moreover, utility rate regulators would allow the utility to charge rates that are appropriate to cover the costs that are in fact incurred.
\textsuperscript{233} A nuclear power facility is a regulated public utility earns a guaranteed level before-tax rate of return that resembles that of an annuity. Economic depreciation of an annuity is de-cellerated depreciation like that shown in Table 7.
\textsuperscript{234} See Schenk and Graetz, note ??? at 718-21 (cash-method-equivalent treatment).
In conclusion, separate accounting for assets and liabilities pretty much eliminates the possibility that tax accounting can come close to mimicking a Simons income (MTM) result. This is especially the case where asset accounting is already distorted. In a second-best world, improving liability accounting rules (relative to an MTM norm) does not necessarily improve overall economic efficiency.\textsuperscript{235}

V. CONCLUSION

If any conclusion can be reached, it is that there is no simple answer to the proper tax treatment of borrowing and liabilities under the income tax. Although the current tax treatment of borrowing is suitable in the case of debt-financed capital expenditures that yield fully-taxed income, the alternative (PID) treatment is suitable for debt-financed expenses and debt-financed investments that yield severely-under-taxed income. In addition, PID treatment of consumer debt is worth considering on tax and non-tax grounds. Finally, it appears that the accrual method of accounting for expenses (and gross income items) is theoretically inferior to the cash method, and that the OID norm is seriously overrated.

More broadly, one may take away from this discussion additional nourishment than what was brought to the table. An income tax advocate might come away with new ideas for improving the income tax. A consumption tax advocate may be reinforced in the belief that the income tax is hopelessly infected by a pervasive and incoherent realization principle. Anybody can acquire, for better or worse, a fuller grasp of implications of the ability-to-pay norm. \textit{Chacon a son gout!}

\textsuperscript{235} Cite theory of the second best. ??
HOMO SACER, HOMOSEXUAL:

SOME THOUGHTS ON WAGING TAX GUERRILLA WARFARE

Anthony C. Infanti

In fact, it is now clear to most that the social and political forces now holding power are beyond simply opposing issues supportive to LGBT people and have now moved to open warfare against all that they hold in contempt, including and especially the LGBT community. It is then little surprise that LGBT communities are experiencing not only unprecedented attacks politically, but have also been living through an unprecedented and sustained increase in anti-LGBT violence.

—National Coalition of Anti-Violence Programs

Questioning the ostensibly unquestionable premises of our way of life is arguably the most urgent of the services we owe our fellow humans and ourselves.

—Zygmunt Bauman

TAX TIME

I always find January depressing. It isn’t the weather that gets me down, although the gray Pittsburgh skies and the frigid temperatures certainly can be trying. No, it’s the constant barrage of mail from banks, mortgage companies, and my employer, all of whom are so thoughtfully providing me with the information that I need to complete my federal income tax return. Given the “ugh!” that is probably reverberating inside your head as you read this, I’m sure that this plaint would sound trite if I weren’t to immediately confess that I’m a “tax geek,” someone who makes his living teaching and writing about the tax laws.

Alas, I find tax time depressing for reasons different from most. To me, tax time

1 Associate Professor of Law, University of Pittsburgh School of Law. I would like to thank Vivian Curran, David Herring, Leandra Lederman, John Parry, and Lu-in Wang for their thoughtful (and helpful!) comments on prior drafts of this essay. I would also like to thank Eliza Hall for her careful cite-checking of this essay. As always, I must thank Hiên for his unflagging love and support during the writing of this essay (read: for putting up with me while I spent—and spend—so much of my time “chewing books” instead of painting the house).


4 Well, they’re not really being thoughtful; they’re just complying with their legal obligation to send us all of these little pieces of paper. I.R.C. §§ 6041(d), 6049(c), 6050H(d), 6051(a) (2006); Treas. Reg. § 1.6041-2(a)(5) (as amended in 2004); id. § 1.6049-6 (as amended in 1999); id. § 1.6050H-2(b) (as amended in 2000); id. § 31.6051-1 (as amended in 2004).

5 See http://www.law.pitt.edu/infanti/cv.htm for a list of the tax courses that I teach and a list of my publications.
is more than the occasion for fulfilling my obligation to defray a portion of the cost of government; it is an annual reminder of my difference—and of my oppression by the government because of that difference. Completing my federal income tax return reminds me that the government has singled out for condemnation my partner and me, my sister and her partner, and every other lesbian and gay man in the United States.

When the W-2s and 1099s begin to appear in my mailbox, I can’t help but think how the federal government legally erased even the possibility of a relationship for me when it enacted the Defense of Marriage Act (DOMA). From the perspective of the federal government, my marriage to my partner in Toronto, Canada, never really happened—and, for that matter, never could happen. Each year, when tax season comes around, I feel the legal eraser scraping against me once again as the federal government returns to ensure that it has removed all trace of my relationship. As if to continually reaffirm its success in wiping away the connection between us (or, perhaps, because it never really can succeed), the federal government forces my partner and me to act as if we were total strangers by demanding that we file two “single” tax returns every April 15.

But even within the diaphanous realm of federal tax law, we cannot truly be made “single.” Because our lives are intertwined financially and emotionally, when the federal government designates us as legal strangers, it can, at most, banish us to that uncomfortable and uncertain space between “single” and “joint.” Life in this tax “limbo” is in some ways more precarious than DOMA’s outright condemnation would seem to indicate. In tax limbo, members of lesbian and gay couples are told what they are not (i.e., married), but they are never told what they are (and, concomitantly, how they should report transactions between them). The existence of this limbo opens the way for the federal government to visit further, more dehumanizing, indignities upon us: It allows the federal government to invade the sanctity of our homes—and of our relationships—to demand that we account for our every move, with our partners and with others, or suffer consequences that range from confiscatory monetary sanctions (i.e., interest, penalties, and, of course, interest on the penalties) to imprisonment.

So, when the dreams of sugar plums in December give way to nightmares about what might better be termed the lesbian and gay circle of tax hell in January, I can’t help but feel haunted by the voices of the reactionary congressmen who enacted DOMA as they repeatedly deprecate my relationship by referring to it as a “marriage”—with the

---

8 See Mark Schwanhausser, Domestic Partners in Tax-Return Limbo, SAN JOSE MERCURY NEWS, Feb. 23, 2006, at 1A (this thought has clearly not occurred to me alone). I find it quite ironic that, even as the Roman Catholic Church moves toward abandoning the notion of limbo, it still manages to persist in the federal tax laws. Ian Fisher, Limbo, an Afterlife Tradition, May Be Doomed by the Vatican, N.Y. TIMES, Dec. 28, 2005, at A1.
10 Cf. DANTE, THE DIVINE COMEDY: HELL, PURGATORY, HEAVEN, at Hell, Canto 11, ll. 50–51 (Peter Dale trans., Anvil Press Poetry 1996) (1472) (indicating that a separate ring of the seventh circle of hell is devoted to “Both Sodom and Cahors, and all who name/God with disparagement within their hearts”).
quotation marks that mark it as a sham, a failed and hopelessly failing attempt at establishing a lasting, loving tie with another human being.  

**HOMO SACER AND THE STATE OF EXCEPTION**

Fortunately, tax time this year has not felt quite so oppressive. I won’t go so far as to say that it’s been cheerful, but I certainly would say that it has been more thoughtful. A book suggested by a (now, unfortunately, former) Pitt Law colleague while we were chatting at the AALS annual meeting provoked me to reflect on this seasonal affliction and to approach it from a different perspective. The book, *Homo Sacer: Sovereign Power and Bare Life* by Giorgio Agamben, deals with biopolitics and the nature of sovereignty, two subjects that, at first blush, probably seem to have only the most tenuous of connections with the tax treatment of same-sex couples. But, if you hang on for just a few pages, I promise that you will begin to see the connection.

Although Agamben’s writing can, at times, be abstruse, the title of his book quite clearly points us to the key, intertwined themes of his discussion: sovereign power and bare life (which, as we will see, is essentially synonymous with the figure of *homo sacer* [sacred man]). To provide needed background, I will briefly consider each of these concepts separately and then discuss the manner in which Agamben brings them together.

For Agamben, the basic paradox of sovereignty is that “the sovereign is, at the same time, outside and inside the juridical order.” In other words, the sovereign—and it’s worth noting that Agamben does not confine his discussion to the conventional notion

---

12 Thank you, John Parry!
13 Each January, the Association of American Law Schools, [http://www.aals.org](http://www.aals.org), holds a conference for law professors. In my limited experience, I seem to learn more from talking with colleagues in the hallways than from attending any of the panel presentations.
14 Tax affective disorder, maybe? See Nat’l Mental Health Assoc., *Seasonal Affective Disorder*, [http://www.nmha.org/infoctr/factsheets/27.cfm](http://www.nmha.org/infoctr/factsheets/27.cfm) (last visited Feb. 17, 2006) (“Some people suffer from symptoms of depression during the winter months, with symptoms subsiding during the spring and summer months. This may be a sign of Seasonal Affective Disorder (SAD). SAD is a mood disorder associated with depression episodes and related to seasonal variations of light.”).

The sovereign exception (as zone of indistinction between nature and right) is the presupposition of the juridical reference in the form of its suspension. Inscribed as a presupposed exception in every rule that orders or forbids something (for example in the rule that forbids homicide) is the pure and unsanctionable figure of the offense that, in the normal case, brings about the rule’s own transgression (in the same example, the killing of a man not as natural violence but as sovereign violence in the state of exception).

*Id.* at 21.
of “sovereign” as “monarch” or “absolute ruler”; for this purpose, it makes no difference what form the sovereign takes\textsuperscript{18}—has the power to decide both what is included in and what is excluded from the juridical order. The sovereign’s power to suspend the juridical order—to declare what Agamben refers to as a “state of exception”—is an example of this power to carve out a sphere that is excluded from the juridical order and, therefore, is not subject to regulation by law.\textsuperscript{19} The Bush Administration’s position with respect to the legal status of the detainees held at Guantanamo Bay, Cuba is a clear example of a sovereign’s declaration of such a state of exception.\textsuperscript{20}

For Agamben, when the sovereign declares a state of exception, what the sovereign excludes from the juridical order is not merely cast out from it. Instead, Agamben posits a complex, continuing relationship between this exception and the general rule.\textsuperscript{21} In his view, “the exception does not subtract itself from the rule; rather the rule, suspending itself, gives rise to the exception.”\textsuperscript{22} Agamben sees the rule as abandoning the exception, but, nonetheless, still “maintaining itself in relation to the exception.”\textsuperscript{23} It is in this way that the rule “first constitutes itself as a rule.”\textsuperscript{24} In fact, Agamben questions whether one can truly say that the exception has been placed outside the juridical order,\textsuperscript{25} because the general rule (i.e., what is included within the juridical

\textsuperscript{18} See id. at 30 (“The principle according to which sovereignty belongs to law, which today seems inseparable from our conception of democracy and the legal State, does not at all eliminate the paradox of sovereignty; indeed it even brings it to the most extreme point of its development.”).

And only because biological life and its needs had become the politically decisive fact is it possible to understand the otherwise incomprehensible rapidity with which twentieth-century parliamentary democracies were able to turn into totalitarian states and with which this century’s totalitarian states were able to be converted, almost without interruption, into parliamentary democracies . . . . [T]he only real question to be decided was which form of organization would be best suited to the task of assuring the care, control, and use of bare life. Once their fundamental referent becomes bare life, traditional political distinctions . . . lose their clarity and intelligibility and enter into a zone of indistinction.

\textsuperscript{19} Agamben traces the history of the state of exception in a more recent book, GIORGIO AGAMBEN, STATE OF EXCEPTION 11–22, 41–51 (Kevin Attell trans., Univ. of Chicago Press 2005) (2003) [hereinafter AGAMBEN, EXCEPTION].


\textsuperscript{21} See AGAMBEN, HOMO SACER, supra note 15, at 17–18 (“[W]hat is excluded . . . is not, on account of being excluded, absolutely without relation to the rule. On the contrary, what is excluded in the exception maintains itself in relation to the rule in the form of the rule’s suspension. The rule applies to the exception in no longer applying, in withdrawing from it.”).

\textsuperscript{22} Id. at 18.

\textsuperscript{23} Id.

\textsuperscript{24} Id.

\textsuperscript{25} Id.

If the exception is the structure of sovereignty, then sovereignty is not an exclusively political concept, an exclusively juridical category, a power external to law (Schmitt), or the supreme rule of the juridical order (Hans Kelsen): it is the originary structure in which law refers to life and includes it in itself by suspending it. Taking up Jean-Luc Nancy’s suggestion, we shall give the name \textit{ban} (from the old Germanic term that
order) only really takes on meaning when it can be compared and contrasted with the exception (i.e., what has been excluded from the juridical order).\(^{26}\) Indeed, Agamben later pointedly states that the “[l]aw is made of nothing but what it manages to capture inside itself through the inclusive exclusion of the exceptio: it nourishes itself on this exception and is a dead letter without it.”\(^{27}\) In elaborating on the content-giving function of the exception, Agamben quotes at length from Carl Schmitt’s discussion of the structure and importance of the exception:

“The exception is more interesting than the regular case. The latter proves nothing; the exception proves everything. The exception does not only confirm the rule; the rule as such lives off the exception alone . . . . ‘The exception explains the general and itself. And when one really wants to study the general, one need only look around for a real exception. It brings everything to light more clearly than the general itself. After a while, one becomes disgusted with the endless talk about the general—there are exceptions. If they cannot be explained, then neither can the general be explained. Usually the difficulty is not noticed, since the general is thought about not with passion but only with comfortable superficiality. The exception, on the other hand, thinks the general with intense passion.’”\(^{28}\)

\[\text{id. at 28–29.}\] Agamben's assertion at the end of the quoted passage that “[t]here is nothing outside the law” (“non c’è un fuori della legge”), GIORGIO AGAMBEN, HOMO SACER: IL POTERE SOVRANO E LA NUDA VITA 34 (1995) (many thanks to my former student, Joe Gulino, for confirming the translation), appears to be a paraphrasing of Jacques Derrida’s famous statement “[t]here is nothing outside of the text [there is no outside-text; il n’y a pas de hors-texte].” JACQUES DERRIDA, OF GRAMMATOLOGY 158 (Gayatri Chakravorty Spivak trans., Johns Hopkins Univ. Press 3d ed. 1997) (1967).

\(^{26}\) See AGAMBEN, HOMO SACER, supra note 15, at 19 (indicating that the sovereign’s decision regarding the exception “is the originary juridico-political structure on the basis of which what is included in the juridical order and what is excluded from it acquire their meaning.”); see also id. at 17 (“Through the state of exception, the sovereign ‘creates and guarantees the situation’ that the law needs for its own validity.”).

\(^{27}\) Id. at 27.

\(^{28}\) Id. at 16 (quoting CARL SCHMITT, POLITISCHE THEOLOGIE: VIER KAPITEL ZUR LEHRE VON DER SOUVERÄNITÄT 19–22 (1922)). Carl Schmitt was one of the foremost Nazi legal theorists, id. at 169, and Agamben particularly draws on his work in discussing the concentration camp as the “pure space of exception.” Id. at 134; see also id. at 169, 171, 175, 184.

It is also worth noting that Agamben’s (and, in turn, Schmitt’s) discussion of the “exception” is very similar to the deconstructionist notion of the “trace,” which is the idea that “the terms in a hierarchical opposition rely for their coherence on the differentiation between them” and, therefore, each bear a “trace” of the other. J.M. Balkin, Deconstructive Practice and Legal Theory, 96 YALE L.J. 743, 752 (1987); see also DERRIDA, supra note 25, at 46–47. For further discussion of the “trace,” see Infanti, supra note 16, at 756–57.
Now let us turn to a brief discussion of “bare life” that will naturally lead us to the connection that Agamben makes between sovereign power and bare life. “Bare life” is exactly what it sounds like: “bare natural life—which is to say, the pure fact of birth.” In this regard, Agamben distinguishes bare life from political life, which is “a qualified life, a particular way of life.”

For Agamben, the Roman legal figure of homo sacer is the quintessential form of bare life. Homo sacer is the sacred man “who may be killed and yet not sacrificed.” To more fully describe the enigmatic figure of homo sacer, Agamben quotes from Pompeius Festus’ treatise, *On the Significance of Words*:

“The sacred man is the one whom the people have judged on account of a crime. It is not permitted to sacrifice this man, yet he who kills him will not be condemned for homicide; in the first tribunitian law, in fact, it is noted that ‘if someone kills the one who is sacred according to the plebiscite, it will not be considered homicide.’ This is why it is customary for a bad or impure man to be called sacred.”

To Agamben, homo sacer represents no more than a bare, nonpolitical life because his killing (by anyone) will not be legally punished as a homicide; yet, at the same time, neither can that killing be perpetrated using the machinery of the law (e.g., through the ritual practice of a trial followed by the imposition of the death penalty).

For precisely the same reason, homo sacer, the quintessential form of bare life, is

---

29 AGAMBEN, HOMO SACER, supra note 15, at 127.
30 Id. at 1.

In the classical world, however, simple natural life is excluded from the polis in the strict sense, and remains confined—as merely reproductive life—to the sphere of the oikos, ‘home.’ . . . At the beginning of the *Politics*, Aristotle takes the greatest care to distinguish the oikonomos (the head of an estate) and the despote (the head of the family), both of whom are concerned with the reproduction and the subsistence of life, from the politician, and he scorns those who think the difference between the two is one of quantity and not of kind.

Id. at 2

One of the essential characteristics of modern biopolitics (which will continue to increase in our century) is its constant need to redefine the threshold in life that distinguishes and separates what is inside from what is outside. Once it crosses over the walls of the oikos and penetrates more and more deeply into the city, the foundation of sovereignty—nonpolitical life—is immediately transformed into a line that must be constantly redrawn.

Id. at 131.
31 Id. at 8 (“[t]he protagonist of this book[,] bare life, that is, the life of homo sacer (sacred man), who may be killed and yet not sacrificed”).
32 Id.
33 Id. at 71. Agamben describes homo sacer as an enigma because of the apparent contradiction in allowing a sacred man to be killed with impunity, so long as he was not “put to death according to ritual practices.”
likewise the embodiment of the state of exception.\textsuperscript{34} \textit{Homo sacer} embodies the state of exception because the juridical order has been suspended with regard to him; again, neither will his killing be punished by the law, nor will it be effected under and through the law. \textit{Homo sacer} has, in effect, been abandoned by the law—placed in a state of exception that gives meaning and content to that which is included within the juridical order.

Tracing the long (though veiled) history of the connection between bare life and the sovereign decision concerning the state of exception,\textsuperscript{35} Agamben argues that the production of bare life actually forms the very heart of sovereign power and is nothing less than the “originary ‘political’ relation.”\textsuperscript{36} According to Agamben, the only thing that is new to our era is the extent to which the state of exception “comes more and more to the foreground as the fundamental political structure and ultimately begins to become the rule,”\textsuperscript{37} rendering all of us virtual \textit{homines sacri} who may at any time be designated by the sovereign as subject to being killed but not sacrificed.\textsuperscript{38}

\textsuperscript{34} See id. at 83 (“\textit{homo sacer} presents the originary figure of life taken into the sovereign ban and preserves the memory of the originary exclusion through which the political dimension was first constituted”); see supra note 25 for a discussion of the sovereign ban.

\textsuperscript{35} Using \textit{homo sacer} as the paradigm, Agamben traces bare life as the originary political relation or element through the Roman \textit{vita necisque potestas} (i.e., “the unconditional authority [\textit{potestà}] of the \textit{pater} over his sons,” AGAMBEN, \textit{HOMO SACER}, supra note 15, at 87), the political body of the king and the figure of the \textit{devotus} (i.e., one “who consecrates his own life to the gods of the underworld in order to save the city from a grave danger,” id. at 96), as well as the figure of the werewolf. \textit{Id.} at 87–111. Continuing into the modern era, Agamben detects the presence of \textit{homo sacer} in the 1679 writ of \textit{habeas corpus} and the French Declaration of the Rights of Man and Citizen in 1789, \textit{Id.} at 123–35, in the status of refugees, \textit{Id.} at 131–34, and in determining the time of death of comatose patients (such as Karen Quinlan), \textit{Id.} at 160–65. Agamben’s exploration of bare life and the state of exception culminates with a discussion of Nazi eugenics and the concentration camp as the “pure space of exception.” \textit{Id.} at 134; see \textit{Id.} at 136–59, 166–80.

\textsuperscript{36} What this work has had to record among its likely conclusions is precisely that the two analyses cannot be separated, and that the inclusion of bare life in the political realm constitutes the original—if concealed—nucleus of sovereign power. \textit{It can even be said that the production of a biopolitical body is the original activity of sovereign power.} In this sense, biopolitics is at least as old as the sovereign exception. Placing biological life at the center of its calculations, the modern State therefore does nothing other than bring to light the secret tie uniting power and bare life, thereby reaffirming the bond . . . between modern power and the most immemorial of the \textit{arcana imperii}.

\textit{Id.} at 6.

If our hypothesis is correct, sacredness is . . . the originary form of the inclusion of bare life in the juridical order, and the syntagm \textit{homo sacer} names something like the originary ‘political’ relation, which is to say, bare life insofar as it operates in an inclusive exclusion as the referent of the sovereign decision.

\textit{Id.} at 85; see also AGAMBEN, \textit{EXCEPTION}, supra note 19, at 87–88 (“Bare life is a product of the machine and not something that preexists it, just as law has no court in nature or in the divine mind.”).

\textsuperscript{37} AGAMBEN, \textit{HOMO SACER}, supra note 15, at 20.

\textsuperscript{38} And if in modernity life is more and more clearly placed at the center of State politics (which now becomes, in Foucault’s terms, biopolitics), if in our age all citizens can be said, in a specific but extremely real sense, to appear virtually as \textit{homines sacri}, this is
THE LESBIAN AND GAY EXCEPTION

Despite the endless “war on terror” and the extraordinary powers arrogated by the President in its name (e.g., warrantless wiretapping, secret CIA detention centers, and the torture perpetrated at the prisons at Guantanamo Bay and Abu Ghraib) as well as the government intervention in the Terri Schiavo case,49 most Americans would probably find it difficult to accept Agamben’s assertion that we now live in a permanent state of exception and are all virtually homines sacri. I imagine, however, that most lesbians and gay men would, like me, readily identify with the figure of homo sacer.40

Like homo sacer, lesbians and gay men are the exception that gives meaning and content to the general rule that is heterosexual privilege. Naturally, heterosexuals encounter difficulties seeing the pervasiveness of their privilege because they live and breathe the general rule. They cannot see the ease with which they speak of husbands and wives or parents and children (all of whom share in common the links of heterosexual reproductive activity) or the brazen way that they celebrate their marriages, openly wear their wedding rings, and display photos of their families at work. It is only when heterosexuals directly confront the lesbian and gay exception41—for example, the very real instance of the lesbian who lost her job for participating in a religious wedding ceremony with her partner42—that heterosexuals have occasion to realize the meaning and extent of their privilege. In contrast, for lesbians and gay men, it seems that heterosexual privilege is both ever-present and inescapable. We cannot help but constantly be faced with the ways in which we are treated as the exception to the heterosexual rule.

In the Internal Revenue Code, lesbians and gay men served for decades as the possible only because the relation of ban has constituted the essential structure of sovereign power from the beginning.

Id. at 111; see also id. at 124 (“modern democracy does not abolish sacred life but rather shatters it and disseminates it into every individual body, making it into what is at stake in political conflict”); id. at 140 (“Bare life is no longer confined to a particular place or a definite category. It now dwells in the biological body of every living being.”).


40 In this regard, I feel constrained to disaffirm the (currently) possible fulfillment of the condition upon which the following statement hinges: “If today there is no longer any one clear figure of the sacred man, it is perhaps because we are all virtually homines sacri.” AGAMBEN, HOMO SACER, supra note 15, at 115.


42 Shahar v. Bowers, 114 F.3d 1097 (11th Cir. 1997). The federal appeals court upheld the decision of the Georgia attorney general to withdraw the job offer on the ground that the attorney general could reasonably conclude that the decision of one of his employees to “marry” another woman would confuse the public, would cause the public to question his office’s credibility, would interfere with his office’s ability to handle the enforcement of the Georgia sodomy law, and would form a reasonable basis for the attorney general to lose confidence in that employee’s “ability to make good judgments as a lawyer for the Law Department.” Id. at 1110.
unspoken foils for heterosexual married couples. Lesbians and gay men were the irreducible singles—the confirmed bachelor or maiden aunt—the singles who would never (and could never) get married. By enacting DOMA, Congress did no more than set the general rule down in writing: “‘marriage’ means only a legal union between one man and one woman as husband and wife.” In an area where the government acquires such sensitive information about individuals that its interactions are shielded from public view, heterosexual married couples are afforded a privileged zone of privacy in which even the Internal Revenue Service (Service) will not enter. But again, this general rule means nothing until it is juxtaposed with the unstated, yet implicit exception: the lesbians and gay men who, shorn of any euphemisms, are now irreducibly and irretrievably single, and whose relationships, far from being afforded privacy, are subject to intense government scrutiny.

And, like homo sacer, each lesbian and gay man has been reduced to a bare life, one that may be killed but not sacrificed. We are no longer put to death by the “sanctioned ritual practices” of the state simply for being homosexual; that barbarity was left behind long ago. But the state’s pervasive marking of us as the exception to the

43 Unless, of course, they were to accede to social pressure and marry someone of the “opposite” sex in order to pass as straight.
44 1 U.S.C. § 7 (2006). As my colleague Vivian Curran pointed out to me, the sociologist Zygmunt Bauman would likely argue that Congress did more than simply set the general rule down in writing; it exerted the only power that it has left in our “globalized” world—the power to repress the weak and vulnerable. (I place the word “globalized” in quotation marks because, as Bauman notes, the more the word “globalization” is used, the less it really seems to signify. BAUMAN, supra note 3, at 1.)

46 See id. §§ 1041, 2056, 2523.
47 See supra note 9 and accompanying text.
48 AGAMBEN, HOMO SACER, supra note 15, at 89.
49 During the colonial period in New England, sodomy was punishable by death. BYRNE FONE, HOMOPHOBIA: A HISTORY 329 (2000). There are records of men being executed as well as records of men being severely whipped, burned with a hot iron, and then made permanent outcasts for engaging in
general rule of heterosexual privilege does open the space for individuals to attack and kill us with impunity. As research psychologist Dr. Gregory Herek has observed:

Whereas psychological heterosexism [i.e., the manifestation of heterosexism in an individual’s actions and attitudes]\(^5\) may not always be the principal reason for an anti-gay attack (e.g., a gang might well have selected another type of “outsider” as a suitable victim), the importance of cultural heterosexism [i.e., the manifestation of heterosexism in societal customs and institutions]\(^6\) cannot be underestimated. For it is cultural heterosexism that defines gay people as suitable targets that can be “used” for meeting a variety of psychological needs. And anti-gay attacks, regardless of the perpetrator’s motivation, reinforce cultural heterosexism. Thus, when a teenage gang member attacks a gay man on the street, it is a hate crime not because hate necessarily was the attacker’s primary motive (it may or may not have been) but because the attack expresses cultural hostility, condemnation, and disgust toward gay people and because it has the effect of terrorizing the individual victim as well as the entire lesbian and gay community. The attack in effect punishes the gay person for daring to be visible.\(^7\)

Although this observation was made more than a decade ago, it could just as aptly have been made today. Anti-gay violence persists at high levels in American society. When adjusted for population size, lesbians and gay men report higher rates of bias crimes than do African-Americans or Jewish people, and they report significantly more crimes against the person than either of those groups.\(^8\) Disturbingly, it appears that anti-gay violence spikes whenever the lesbian and gay community finds itself in the spotlight.

Consider, by way of example, the years 2003 and 2004: In the geographic area covered by the National Coalition of Anti-Violence Programs (which includes less than

---


\(^7\) Herek, supra note 50, at 164; see also Karen Franklin, Unassuming Motivations: Contextualizing the Narratives of Antigay Assailants, in STIGMA AND SEXUAL ORIENTATION: UNDERSTANDING PREJUDICE AGAINST LESBIANS, GAY MEN, AND BISEXUALS 1, 20 (Gregory M. Herek ed., 1998) (“antigay violence can be seen primarily as an extreme manifestation of pervasive cultural norms rather than as a manifestation of individual hatred”).

\(^8\) William B. Rubenstein, The Real Story of U.S. Hate Crimes Statistics: An Empirical Analysis, 78 TUL. L. REV. 1213, 1232 (2004). It is worth noting that Rubenstein’s analysis uses the bias crime statistics compiled by the Federal Bureau of Investigation. As described infra note 54, these statistics only reflect reported bias crimes and therefore underestimate the total number of bias crimes that occur in the United States each year.
30% of the national population),\textsuperscript{54} the number of incidents of violence against gay men, lesbians, bisexuals, and transgender individuals increased 8% from 2002 to 2003\textsuperscript{55} and increased an additional 4% from 2003 to 2004.\textsuperscript{56} Although there was a 4% decrease in the number of victims suffering injuries in 2003, the number of victims suffering serious injuries rose 3%,\textsuperscript{57} and the number of murder victims rose 80% (from 10 in 2002 to 18 in 2003).\textsuperscript{58} Again in 2004, despite a 2% decrease in the number of victims suffering injuries,\textsuperscript{59} the number of victims suffering serious injuries rose an astounding 20%\textsuperscript{60} and the number of murder victims rose 11% (to 20 in 2004).\textsuperscript{61} Providing support for the existence of a “spotlight” effect, there was a noticeable spike in anti-gay violence in the latter half of 2003, when the decisions in \textit{Lawrence v. Texas} (striking down Texas’ sodomy statute) and \textit{Goodridge v. Department of Public Health} (extending the right to marry to same-sex couples in Massachusetts) were issued, and that spike continued into the first half of 2004.\textsuperscript{62}

The ability to physically menace and even kill lesbians and gay men with impunity stems in part from the fact that these crimes often go unreported—out of fear of further harassment from the police.\textsuperscript{63} Even when these crimes are reported, advocacy groups find it necessary to press for the investigation of complaints.\textsuperscript{64} According to the

\textsuperscript{54} NCAVP 2004 REPORT, supra note 2, at 18 (stating that the report covers “approximately 27% of the nation’s population") (citation omitted); NAT’L COALITION OF ANTI-VIOLENCE PROGRAMS, ANTI-LESBIAN, GAY, BISEXUAL AND TRANSGENDER VIOLENCE IN 2003, at 18–19 (2004), http://www.ncavp.org/common/document_files/Reports/2003NCAVP_HV_Report.pdf (stating that the report covers “29.3% of the nation’s population") [hereinafter NCAVP 2003 REPORT]. The Federal Bureau of Investigation (FBI) also reports bias crime statistics, including statistics with respect to crimes that are motivated by sexual orientation bias; however, these reports significantly underreport the level of anti-gay violence in the United States. The FBI report for 2003, which covers a geographic area including nearly 83% of the national population, only reported 1,239 incidents of violence motivated by sexual orientation bias, which is far below that reported by the National Coalition of Anti-Violence Programs with respect to a far smaller portion of the national population. FED. BUREAU OF INVESTIGATION, U.S. DEP’T OF JUSTICE, HATE CRIME STATISTICS: 2003, at 1, 9 (2004), available at http://www.fbi.gov/ucr/03hc.pdf; see also NCAVP 2003 REPORT, supra, at 18–19. The FBI’s underreporting of sexual orientation-motivated bias crimes has been attributed to a number of factors, including the victim’s desire not to be outed and lesbians’ and gay men’s general distrust of the police due to a history of harassment at their hands. DONALD ALTSCHILLER, HATE CRIMES: A REFERENCE HANDBOOK 27–28 (2d ed. 2005).

\textsuperscript{55} NCAVP 2003 REPORT, supra note 54, at 21. This represented a reversal of a general downward trend in anti-gay violence over the previous five-year period, which was part of the general decrease in crime nationally—although anti-gay violence “did not fall as far or as rapidly as violent crime in general” during that period. Id. at 16.

\textsuperscript{56} NCAVP 2004 REPORT, supra note 2, at 24.

\textsuperscript{57} NCAVP 2003 REPORT, supra note 54, at 4.

\textsuperscript{58} Id. at 21.

\textsuperscript{59} NCAVP 2004 REPORT, supra note 2, at 27.

\textsuperscript{60} Id.

\textsuperscript{61} Id. at 25.

\textsuperscript{62} Id. at 16; NCAVP 2003 REPORT, supra note 54, at 15. This paralleled earlier spikes: first, in New York City in June 1994, when that city hosted both the Gay Games and the Stonewall 25 celebration, and then nationally in March and April 1997, when Ellen DeGeneres and her character on her eponymous television show simultaneously came out of the closet. NCAVP 2003 REPORT, supra note 54, at 14.

\textsuperscript{63} NCAVP 2004 REPORT, supra note 2, at 43.

\textsuperscript{64} Id. (“Often in the experience of NCAVP members, even victims of brutal anti-LGBT assaults will hesitate to file police reports, and for those who do, a good portion of the services that NCAVP agencies provide is concerned with persuading police to act on their complaints in a meaningful way.”).
National Coalition of Anti-Violence Programs, in 2004, there was an 82% increase in the number of bias crime complaints that police refused to take, and a stunning 66% of the bias crime complaints that were taken resulted in no arrest.\footnote{Id. at 44.} It should come as little surprise then to hear reports of reactionaries calling for an “open season” on lesbians and gay men.\footnote{Bob Hague, \textit{Voicemail Message Suggests “Open Season” on Gays} (Wis. Radio Network Dec. 6, 2005), \url{http://www.wrn.com/gestalt/go.cfm?objectid=E78DA9DB-FA31-41FE-835BAA01956B977&dbtranslator=local.cfm} (containing a story about, as well as a link to, a voicemail message that called for an “open season” on lesbians and gay men; the message had been left for a state legislator on the eve of a debate about same-sex marriage); \textit{see also} Boyd County High Sch. Gay Straight Alliance v. Bd. of Educ., 258 F. Supp. 2d 667, 670 n.1 (E.D. Ky. 2003) (“One example of the harassment includes students in . . . English class stating that they needed to take all the fucking faggots out in the back woods and kill them.”); \textit{AM. CIVIL LIBERTIES UNION, 2006 WORKPLAN 6} (2006) (on file with author) (in speaking of the same school, stating that “[t]he anti-harassment training part of the settlement was important because there was widespread anti-gay harassment in the school, and teachers and administrators were doing little to stop it. For example, the school’s Model United Nations once adopted a resolution declaring an ‘open hunting season’ on gay students.”).}

**OF PARABLES AND PROVOCATEURS**

As homines sacri living in a virtual state of exception,\footnote{On the notion of a “virtual” state of exception, see \textit{AGAMBEN, EXCEPTION, supra} note 19, at 3 (in this translation, the terminology “real” versus “fictitious” state of exception is used in place of the terminology “real” versus “virtual” state of exception, which is used in \textit{AGAMBEN, HOMO SACER, supra} note 15) and at 4 (“Though the notions of \textit{state of siege} and \textit{martial law} express a connection with the state of war that has been historically decisive and is present to this day, they nevertheless prove to be inadequate to define the proper structure of the phenomenon, and they must therefore be qualified as \textit{political} or \textit{fictitious} . . . . The history of the term \textit{fictitious} or \textit{political state of siege} is instructive . . . . It goes back to the French doctrine that . . . provided for the possibility of a state of siege that the emperor could declare whether or not a city was actually under attack or directly threatened by enemy forces . . . .”)).} lesbians and gay men should be particularly interested in Agamben’s interpretation of Franz Kafka’s parable \textit{Before the Law}, which Kafka later incorporated into chapter nine of his book \textit{The Trial}. It is worth reproducing this short parable in full before considering Agamben’s interpretation of it:

[B]efore the Law stands a doorkeeper. To this doorkeeper there comes a man from the country who begs for admittance to the Law. But the doorkeeper says that he cannot admit the man at the moment. The man, on reflection, asks if he will be allowed, then, to enter later. “It is possible,” answers the doorkeeper, “but not at this moment.” Since the door leading into the Law stands open as usual and the doorkeeper steps to one side, the man bends down to peer through the entrance. When the doorkeeper sees that, he laughs and says: “If you are so strongly tempted, try to get in without my permission. But note that I am powerful. And I am only the lowest doorkeeper. From hall to hall, keepers stand at every door, one more powerful than the other. And the sight of the third man is already more than even I can stand.” These are difficulties which the man from the country has not expected to meet, the Law, he thinks, should be accessible.
to every man and at all times, but when he looks more closely at the
doorkeeper in his furred robe, with his huge, pointed nose and long, thin,
Tartar beard, he decides that he had better wait until he gets permission to
enter. The doorkeeper gives him a stool and lets him sit down at the side
of the door. There he sits waiting for days and years. He makes many
attempts to be allowed in and wears the doorkeeper with his importunity.
The doorkeeper often engages him in brief conversation, asking him about
his home and about other matters, but the questions are put quite
impersonally, as great men put questions, and always conclude with the
statement that the man cannot be allowed to enter yet. The man, who has
equipped himself with many things for his journey, parts with all he has,
however valuable, in the hope of bribing the doorkeeper. The doorkeeper
accepts it all, saying, however, as he takes each gift: “I take this only to
keep you from feeling that you have left something undone.” During all
these long years the man watches the doorkeeper almost incessantly. He
forgets about the other doorkeepers, and this one seems to him the only
barrier between himself and the Law. In the first years he curses his evil
fate aloud; later, as he grows old, he only mutters to himself. He grows
childish, and since in his prolonged study of the doorkeeper he has learned
to know even the fleas in his fur collar, he begs the very fleas to help him
and to persuade the doorkeeper to change his mind. Finally his eyes grow
dim and he does not know whether the world is really darkening around
him or whether his eyes are only deceiving him. But in the darkness he
can now perceive a radiance that streams inextinguishably from the door
of the Law. Now his life is drawing to a close. Before he dies, all that he
has experienced during the whole time of his sojourn condenses in his
mind into one question, which he has never yet put to the doorkeeper. He
beckons the doorkeeper, since he can no longer raise his stiffening body.
The doorkeeper has to bend far down to hear him, for the difference in
size between them has increased very much to the man’s disadvantage.
“What do you want to know now?” asks the doorkeeper, “you are
insatiable.” “Everyone strives to attain the Law,” answers the man, “how
does it come about, then, that in all these years no one has come seeking
admittance but me?” The doorkeeper perceives that the man is nearing his
end and his hearing is failing, so he bellows in his ear: “No one but you
could gain admittance through this door, since this door was intended for
you. I am now going to shut it.”

For Agamben, the man from the country in this parable is akin to homo sacer; he
is living in a virtual state of exception. The law can be seen as holding the man from the
country in a relation of exception. The “law applies to him in no longer applying, and
holds him in its ban in abandoning him outside itself. The open door destined only for
him includes him in excluding him and excludes him in including him.”

---

69 AGAMBEN, HOMO SACER, supra note 15, at 50.
70 Id.
the state of exception, the law’s abandonment of the man from the country only makes its effect upon him all the more powerful:

For life under a law that is in force without signifying resembles life in the state of exception, in which the most innocent gesture or the smallest forgetfulness can have most extreme consequences. And it is exactly this kind of life that Kafka describes, in which law is all the more pervasive for its total lack of content, and in which a distracted knock on the door can mark the start of uncontrollable trials. . . . [S]o in Kafka’s village the empty potentiality of law is so much in force as to become indistinguishable from life.71

Agamben describes the real danger that faces each of us in this state of exception that becomes indistinguishable from (consumes?) our life as the possibility that we might find ourselves “condemned to infinite negotiations with the doorkeeper or, even worse, that [we] might end by [ourselves] assuming the role of the doorkeeper who, without really blocking the entry, shelters the Nothing onto which the door opens.”72

Probably the most striking aspect of Agamben’s analysis of Kafka’s parable, however, is his treatment of its ending. Agamben does not see in this ending “the irremediable failure or defeat of the man from the country before the impossible task imposed upon him by the Law.”73 When the doorkeeper closes the door that was open only to the man from the country, Agamben “imagine[s] that all the behavior of the man from the country is nothing other than a complicated and patient strategy to have the door closed in order to interrupt the Law’s being in force.”74 Thus, instead of failure, Agamben sees success, “even if [the man from the country] may have risked his life in the process (the story does not say that he is actually dead but only that he is ‘close to the end’).”75 Recapitulating the end of the parable in messianic terms, Agamben contends that:

If one gives the name ‘provocation’ to the strategy that compels the potentiality of Law to translate itself into actuality, then his is a paradoxical form of provocation, the only form adequate to a law that is in force without signifying and a door that allows no one to enter on account

---

71 Id. at 52–53 (emphasis added).
72 Id. at 54.
73 Id. at 55. In this regard, Agamben distinguishes his interpretation from that of Derrida, who sees in this ending the defeat of the man from the country:

At the moment when the man comes to his end, just before his death, the doorkeeper points out to him that he will not reach his destination, or that it will not reach him. The man comes to his end without reaching his end. The entrance is destined for and awaits him alone; he arrives there but cannot arrive at entering; he cannot arrive at arriving. Thus runs the account of an event which arrives at not arriving, which manages not to happen.

JACQUES DERRIDA, Before the Law, in ACTS OF LITERATURE 181, 210 (Derek Attridge ed., 1992) (footnote omitted); see AGAMBEN, HOMO SACER, supra note 15, at 57 (distinguishing Derrida’s interpretation).
74 AGAMBEN, HOMO SACER, supra note 15, at 55.
75 Id.
of being too open. The messianic task of the man from the country . . . might then be precisely that of making the virtual state of exception real, of compelling the doorkeeper to close the door of the Law . . . . For the Messiah will be able to enter only after the door is closed, which is to say, after the Law’s being in force without significance is at an end . . . . [T]he messianic aporias of the man from the country express exactly the difficulties that our age must confront in attempting to master the sovereign ban.\footnote{Id. at 56–57.}

Finally, in closing his discussion of the parable, Agamben makes clear that he does not advocate surrender to the power of the state of exception, but, following the example set by the man from the country, urges resistance to it and, ultimately, the subversion of it.\footnote{Id. at 57–58.}

**TAX GUERRILLA WARFARE**

Agamben’s view of sovereignty—as founded upon force and not upon a “contract or convention”\footnote{Id. at 109.}—should occasion lesbians and gay men to rethink their relationship with the law. It should cause us to question our past—albeit quite natural—privileging of legal approaches for achieving social change.\footnote{E.g., ELLEN ANN ANDERSEN, OUT OF THE CLOSETS & INTO THE COURTS 3, 17–26 (2005); Jules Lobel, Courts as Forums for Protest, 52 UCLA L. REV. 477, 479 (2004) (“The reform upsurge of the 1960s and 1970s witnessed a transformation in the role of the judiciary, particularly the federal judiciary. Courts were now often viewed not merely as forums to settle private disputes, but as instruments of societal change.”);}

The violence that Benjamin defines as divine is instead situated in a zone in which it is no longer possible to distinguish between exception and rule. It stands in the same relation to sovereign violence as the state of actual exception, in the eighth thesis, does to the state of virtual exception. This is why (that is insofar as divine violence is not one kind of violence among others but only the dissolution of the link between violence and law) Benjamin can say that divine violence neither posits nor conserves violence, but deposes it. Divine violence shows the connection between the two violences—and even more, between violence and law—to be the single real content of law.

\footnote{Id. at 65.}
For many of us who have suffered oppression or discrimination in any form it is easy to understand the attraction of rights-based approaches. Civil rights initiatives have an immediate, concrete appeal. They promise to secure the basic constitutional rights that lesbians and gay men have previously lived without: freedom from discrimination in areas such as housing, employment, child custody, military service, legal marriage, and spousal benefits. For individuals who live in a country that ostensibly provides these protections to all of its citizens, yet in practice denies them to particular groups, the simple granting of such rights often seems like the ultimate luxury: all we can hope for and, at the same time, too much to hope for.80

While alluring, this legal approach has proved to be both paralyzing and assimilationist. Too often, we do no more than sit idly outside the already open door to the law. We dare not approach the doorkeeper until we receive word from the legal “experts” that the “right” or “best” case with the most “sympathetic” plaintiff has arrived.81 These experts actively discourage anyone who does not fit this ambiguous paradigm—which serves as a repository for every (real or imagined) heterosexual expectation of our non-threatening domestication—82 from even considering the possibility of approaching the doorkeeper.83 When an ostensibly “good” case finally does

William B. Rubenstein, Divided We Litigate: Addressing Disputes Among Group Members and Lawyers in Civil Rights Campaigns, 106 YALE L.J. 1623, 1632 (1997) (describing impact litigation and test cases as cases that “are brought with the intention of establishing a legal precedent that will improve a group’s social situation’’). This privileging is implicit in the ubiquitous references to the “lesbian and gay rights movement.’’ The alternative, more general phrase “lesbian and gay movement” does not limit collective action solely to the attainment of “rights.”

80 DIANE HELENE MILLER, FREEDOM TO DIFFER: THE SHAPING OF THE GAY AND LESBIAN STRUGGLE FOR CIVIL RIGHTS 140 (1998); see also PATRICIA A. CAIN, RAINBOW RIGHTS: THE ROLE OF LAWYERS AND COURTS IN THE LESBIAN AND GAY CIVIL RIGHTS MOVEMENT 1 (2000) (“The role of the lawyers, the legal arguments they construct, and the fine-tuning of these arguments in response to judicial opinions is a central part of any civil rights movement . . . Whether one believes that courts do in fact cause social change, courts are nonetheless crucial in any battle over equal rights.”).

81 Rubenstein, supra note 79, at 1656 (citing as “one of the central complaints about social movements: overreliance on ‘experts’ (lawyers’’); see also, e.g., ANDERSEN, supra note 79, at 85–86 (discussing the choice between pursuing Bowers v. Hardwick or Baker v. Wade in terms of who would be a more “sympathetic” plaintiff); id. at 128–29 (similar discussion with regard to Lawrence v. Texas); id. at 186–87 (gay rights organizations issued a pamphlet discouraging individuals from filing same-sex marriage lawsuits in the wake of Baker v. Vermont, except in “‘the best cases in the best places at the best times’’”).

82 See Devon W. Carbado, Black Rights, Gay Rights, Civil Rights, 47 UCLA L. REV. 1467, 1505–06 & n.148 (“Part of a civil rights strategy involves selecting the ‘right’ (read: most palatable) plaintiffs.”), 1505–17 (exploring gay rights advocates’ choice to use the stories of white lesbians and gay men in challenging the military’s anti-gay policies while they ignored the story of a black man who was the first to mount a successful challenge to those policies); cf. Suzanne B. Goldberg, On Making Anti-Essentialist and Social Constructionist Arguments in Court, 81 OR. L. REV. 629, 661 n.117 (2002) (indicating that during her time at Lambda Legal Defense and Education Fund an effort was made in challenges to sodomy laws and anti-gay measures to obtain a diverse group of plaintiffs).

83 For example, Gay and Lesbian Advocates and Defenders discouraged married same-sex couples from Massachusetts from filing joint tax returns in order to fight the discrimination against them in the federal
happen by, the experts stand and bicker over whether this is truly the right case and the right time to approach the doorkeeper to plead for entrance through the already open door to the law. In the meantime, we sit by suffering needlessly. And pity the poor troublemaker who, on a rare occasion, challenges the experts’ judgment and moves to plead her case directly to the doorkeeper; this rogue soon finds the experts attempting to intercept and block her approach for fear that the doorkeeper will bar the way to all. Thus, it seems that, despite (because of?) our best efforts, we have realized precisely the danger that Agamben warned of: Rather than devoting our energy to fight our oppression under the law, most of us stand immobilized while the “experts” take over “the role of the doorkeeper who, without really blocking the entry, shelters the Nothing onto which the door opens.”

Predictably, this approach has produced decidedly mixed results. It is in the tax laws, because it was not “looking to pick a fight with the IRS.” Gay Newlyweds in Massachusetts Tackle Taxes, ADVOCATE, Jan. 26, 2005, http://www.advocate.com/news_detail_ektid03009.asp.

84 See Rubenstein, supra note 79, at 1635–44 (providing examples of how lesbian and gay civil rights litigation has divided the community within itself (over the appropriate goals of such litigation) and divided the lawyers who are prosecuting this litigation (over the methods for achieving these goals)).

85 For example, several lesbian and gay rights organizations “are doing whatever they can to stop” a lawsuit brought by an attorney who is arguing that the California prohibition against same-sex marriages violates the federal constitution—because his arguments depart from their state-by-state strategy for challenging prohibitions against same-sex marriage. Wyatt Buchanan, Going for Broke in Battle over Gay Vows, S.F. CHRON., Jan. 23, 2006, at A1; see also Rubenstein, supra note 79, at 1632–33 (drawing a distinction between, on the one hand, “professional” public interest litigators (e.g., Lambda Legal Defense and Education Fund, the American Civil Liberties Union’s Lesbian and Gay Rights Project, the National Center for Lesbian Rights, and Gay and Lesbian Advocates and Defenders) and “occasional” civil rights lawyers, on the other), 1680 (suggesting “modifications in the Model Rules concerning the scope of representation, competence, and client loyalty [that] would loosen the individualist hold on professional ethics and enhance a vision of expertise-based professionalism”).

86 AGAMBEN, HOMO SACER, supra note 15, at 54; see also id. at 133 (“In the final analysis, however, humanitarian organizations—which today are more and more supported by international commissions—can only grasp human life in the figure of bare or sacred life, and therefore, despite themselves, maintain a secret solidarity with the very powers they ought to fight.”). Rubenstein claims that “professional civil rights attorneys are often the only attorneys who are actually appointed by and answerable to their communities,” and are, therefore, “less, not more, likely to undermine the political processes of those communities.” Rubenstein, supra note 79, at 1668. Interestingly, Urvashi Vaid has similarly criticized small, homogeneous groups of elites for being unrepresentative, self-appointed spokespersons for the movement, asserting, in contrast, that the Human Rights Campaign, the National Gay and Lesbian Task Force, and similar organizations are “accountable” to the lesbian and gay community. URVASHI VAID, VIRTUAL EQUALITY: THE MAINSTREAMING OF GAY AND LESBIAN LIBERATION 213–18 (1995). Just a few pages later, however, Vaid describes how the Human Rights Campaign, the National Gay and Lesbian Task Force, and similar organizations are undemocratic; elitist; have memberships that comprise only a small fraction of the lesbian and gay community; and, in their governance, are more responsive to the demands of fundraising than to democracy. Id. at 219–23.

87 On the positive side: In 2003, we witnessed the Supreme Court’s stunning reversal of its relatively recent decision in Bowers v. Hardwick, 478 U.S. 186 (1986), when it struck down criminal prohibitions against sodomy on federal constitutional grounds in Lawrence v. Texas, 539 U.S. 558 (2003). Then, not even a year later, in Goodridge v. Department of Public Health, 798 N.E.2d 941 (Mass. 2003), see also Opinions of the Justices to the Senate, 802 N.E.2d 565 (2004), the Massachusetts Supreme Judicial Court extended the right to marry to same-sex couples in Massachusetts on state constitutional grounds. Several other states have enacted domestic partnership or civil union regimes that provide a measure of legal recognition to lesbian and gay couples, including California, Connecticut, Hawaii, Maine, New Jersey, and Vermont. Act Concerning Civil Unions, 2005 Conn. Legis. Serv. P.A. 05-10 (S.S.B. 963) (West); Act to Promote the
interest of the heterosexual majority—or, at the very least, of those heterosexuals who wield power and wish to maintain and enhance that power—to keep lesbians and gay men continually occupied with the task of gaining entrance through the already open door to the law. When the heterosexual majority from time to time accepts the arguments that we proffer, it encourages us to continue with our strategy of constant self-vigilance, which is interrupted only by the occasional cloying supplication for a peek at the light emanating from the open door to the law. We are afforded enough success to keep us engaged, but never enough to allow us to be truly successful. The doorkeeper employed a similar tactic to, at least from his perspective, keep the man from the country continually occupied with the task of gaining entrance through the already open door to the law:

The doorkeeper often engages [the man from the country] in brief conversation, asking him about his home and about other matters, but the questions are put quite impersonally, as great men put questions, and always conclude with the statement that the man cannot be allowed to enter yet. The man, who has equipped himself with many things for his journey, parts with all he has, however valuable, in the hope of bribing the doorkeeper. The doorkeeper accepts it all, saying, however, as he takes each gift: “I take this only to keep you from feeling that you have left something undone.”  

By engaging the law (and, by extension, the heterosexual majority who created it) on its own terms, we only serve to affirm its power over us and legitimize its treatment of us:


On the negative side: These high profile legal successes have been matched by equally high-profile failures. In response to a 1993 decision of the Hawaii Supreme Court that, for the first time, raised the specter of legalized same-sex marriage in the United States, Congress passed, and President Clinton signed into law, the Defense of Marriage Act, Pub. L. No. 104-199, 110 Stat. 2419 (1996). All but a small handful of states have enacted statutory, constitutional, or statutory and constitutional prohibitions against same-sex marriage. NAT’L GAY & LESBIAN TASK FORCE, ANTI-GAY MARRIAGE MEASURES IN THE U.S. AS OF NOVEMBER 15, 2005, http://www.thetaskforce.org/downloads/marriagemap.pdf. Thirteen of the constitutional amendments prohibiting same-sex marriage were approved in 2004—eleven of them by wide margins, ranging from 57% to 86% voting in favor. Alan Cooperman, Same-Sex Bans Fuel Conservative Agenda, WASH. POST, Nov. 4, 2004, at A39; Michael Kranish, Gay Marriage Bans Passed; Measures OK’d in All States Where Eyed, BOSTON GLOBE, Nov. 3, 2004, at A22.

Naturally, these illustrations are no more than that; they are not meant to provide an exhaustive list of the movement’s legal successes or its failures. It is also worth noting that the tentative nature of our legal progress is only underscored when we consider our successes and failures in light of advances on the international level. When our progress is viewed from this wider perspective, it becomes clear that the United States is far from being a leader (and, in fact, is only slowly becoming a follower) in recognizing and remedying lesbian and gay rights issues. Infanti, Tax Protest, supra, at 44–49.


88 KAFKA, supra note 68, at 213–14.
It is almost as if, starting from a certain point, every decisive political event were double-sided: the spaces, the liberties, and the rights won by individuals in their conflicts with central powers always simultaneously prepared a tacit but increasing inscription of individuals’ lives within the state order, thus offering a new and more dreadful foundation for the very sovereign power from which they wanted to liberate themselves.89

And the heterosexual majority derives a tangible benefit from keeping lesbians and gay men in thrall to the law. As *hominis sacri*, the embodiment of the virtual state of exception, lesbians and gay men give content and meaning to the general rule of heterosexual privilege. In fact, to paraphrase Agamben’s astute observation, heterosexual privilege “nourishes itself on this exception and is a dead letter without it.”90

Let us return for a moment to the recurring example of same-sex marriage to illustrate the point. Once the relationship between heterosexuals and homosexuals is posited as one of general rule and exception, it is possible to understand why some heterosexuals might view same-sex marriage as a threat to “traditional” marriage. If same-sex marriage were put on truly equal footing with “traditional” marriage, “traditional” marriage would become a “dead letter.” There would no longer be “traditional” marriage and same-sex “marriage”; there would just be marriage—without quotation marks or qualifiers. In the eyes of some heterosexuals, “traditional” marriage might then be sapped of meaning, because it would lack an authentic exception91 that could serve as a point of distinction or opposition and that, therefore, could give content and meaning to the privileges that currently attend their presumed exclusive access to full marital status.92

So, it is not in the interest of the heterosexual majority (or, again, at the very least, of those heterosexuals who wield power and wish to maintain and enhance that power) simply to let lesbian and gay men pass freely through the already open door to the law. It is, however, in their interest occasionally to reassure lesbians and gay men that efforts to pass through that door are not made in vain—in order that heterosexuals might maintain

---

89 *AGAMBEN, HOMO SACER*, supra note 15, at 121.
90 *Id.* at 27; *see also supra* note 27.
91 It is true that the unmarried (straight and gay) would continue to provide a point of distinction that might be seen as giving meaning to the privileges attendant to marriage. However, it might just as well be said that the benefits of marital status become diluted and marital privilege becomes indistinct once everyone is afforded full and equal access to marriage. Given the strong social norm of monogamous coupling within marriage, the difference between married and unmarried is not so much one of rule and exception, but is more akin to evolutionary stages in social development—in other words, it is socially expected that every eligible single will at some point find a mate and marry. Thus, the line that has implicitly given meaning—and that, for more than a decade now, has *explicitly* given meaning—to marital privilege is not the line between married and unmarried, but the line between those who *can* marry (whether, at any moment, they actually happen to be married or not) and those who *cannot* marry.
92 Through no fault of its own, Massachusetts, which is currently the only state that legally recognizes same-sex marriage, has created a less than fully equal marital status for same-sex couples. Due to the enactment of DOMA, the federal government treats these marriages as a nullity for purposes of federal law, 1 U.S.C. § 7 (2006), and permits other states to (and most, in fact, do) refuse to recognize these marriages, 28 U.S.C. § 1738C (2006). *See supra* note 87. Thus, at best, Massachusetts has created an ephemeral status that has a habit of (not so magically) disappearing—sometimes when couples cross the state borders and other times even while they are still within the state borders (e.g., when federal law is exerting its force on same-sex couples).
lesbians and gay men in a virtual state of exception that gives content and meaning to their own pervasive privileges.

Perhaps, then, Agamben is correct in arguing that true change is possible only when we resist being co-opted into serving as our own doorkeeper, when we cease asking politely for entrance through a door that is already open, and when we instead turn our energies to provoking the closure of that door and the creation of a real state of exception. On more than one occasion, it has been suggested to me that government silence on the tax treatment of same-sex couples is preferable to the message that would be sent should the government choose to speak. But, is it really? Might it not be preferable to force the government to express in words the precise nature and the full extent of its anti-gay animus rather than allowing that animus to remain the unseen and unacknowledged (only by heterosexuals, of course) subtext of lesbian and gay lives?

In other words, the source of our great sorrow and despair—our oppression—may actually be the source of our greatest strength. By provoking the government to close the door of the law on us, we may be able to draw attention to our plight in a way that serves as a catalyst for change. This action and reaction should serve as a call to action for all lesbians and gay men as well as for any potentially sympathetic straight men and women (all of whom should be horrified when they finally realize the true extent and nature of our oppression at the hands of those among them who wield power by exploiting fear and division). At the same time, the threat of social unrest should shock the remainder of straight society out of its complacency.

Conversely, from this perspective, the paralyzing fear of defeat—of being turned away by the doorkeeper—has been the source of our greatest weakness. Our fear has debilitated and domesticated us by turning us into our own doorkeeper:

As the evangelical warning cited by Origen concerning the interpretation of Scripture has it: “Woe to you, men of the Law, for you have taken away the key to knowledge: you yourselves have not entered, and you have not let the others who approached enter either” (which ought to be reformulated as follows: “Woe to you, who have not wanted to enter into the door of the Law but have not permitted it to be closed either”).

Thus, in place of our current approach of engaging the law on its own terms, we might consider using the law strategically in an effort to provoke the closure of the door to the law and, concomitantly, to destabilize heterosexual privilege. We could begin by

---

93 Pat Cain, a tax professor and now Vice Provost at the University of Iowa, has been the recipient of similar suggestions. Patricia Cain, Relitigating Seaborn: Taxing the Community Income of California Registered Domestic Partners 2 (Univ. Iowa, Legal Studies Research Paper No. 05-39, 2006), available at http://ssrn.com/abstract=881763.

94 Cf. AGAMBEN, HOMO SACER, supra note 15, at 133 (“It takes only a glance at the recent publicity campaigns to gather funds for refugees from Rwanda to realize that here human life is exclusively considered (and there are certainly good reasons for this) as sacred life—which is to say, as life that can be killed but not sacrificed—and that only as such is it made into the object of aid and protection. The ‘imploring eyes’ of the Rwandan child, whose photograph is shown to obtain money but who ‘is now becoming more and more difficult to find alive,’ may well be the most telling contemporary cipher of the bare life that humanitarian organizations, in perfect symmetry with state power, need.”).

95 Id. at 54.
recognizing that, despite the quotation marks that so frequently surround the phrase, the “culture war” is more than just a rhetorical device for the reactionary right. It is a very real war, and lesbians and gay men have too often found themselves the victims of reactionary violence. Clearly outnumbered by our foes, we might take a page from the government’s playbook when it litigates against its citizens, and adopt (and adapt) the tactics of guerrilla warfare—using the law to harass the government and provoke it to close the door to the law firmly against us—in an attempt to erode support for the hetero status quo among the “civilian” population. I realize that this probably sounds like quite a radical suggestion; yet, as we will see, it might require only the most ordinary of action to accomplish.

Given the title of this essay and my self-professed status as a tax geek, I’m sure that you won’t be surprised to learn that I think that tax would be the perfect area in which to test these guerrilla warfare tactics against the government. Despite being an area of the law that touches the life of nearly every lesbian and gay man, tax is the one door to the law that is generally left unattended by the experts who have been co-opted into serving as doorkeepers. As a result, tax is one area of the law where we will be able to approach the doorkeeper of the law directly, without having to pass through a gauntlet of experts attempting to dissuade us—or, worse, actively prevent us—from making our way to the already open door to the law. Approaching such a relatively unguarded door to the law may provide us the advantage of surprise in our attack (depending, of course, on who reads this essay).

In addition to being an unguarded approach, tax has particular attributes that are well-suited to its use in waging guerrilla warfare. Guerrilla warfare is associated with “small, mobile and flexible combat groups” that engage in “long, low-intensity confrontation” to “destabilize an authority.” In contrast to conventional civil rights litigation, which normally involves a single or select group of plaintiffs who file suit on behalf of (even if not in the name of) a larger class or group of individuals, tax lends itself to a more diffuse approach. Although taxpayers have, on occasion, filed class action lawsuits, they normally interact with the Service and the courts on an individual basis,

96 See supra text accompanying notes 2 and 54–66.
97 Dunn v. Comm’r, 301 F.3d 339, 349 (5th Cir. 2002) (describing the IRS’ tactics in litigation concerning the valuation of a block of stock as “guerilla warfare”).
98 There was a time when critical legal scholars were accused of engaging in guerrilla warfare. Guyora Binder, On Critical Legal Studies as Guerrilla Warfare, 76 GEO. L.J. 1, 1 n.4 (1987). This analogy was rejected, however, as being a less than accurate description of critical legal studies. See id. at 8–13; David Fraser, If I Had a Rocket Launcher: Critical Legal Studies as Moral Terrorism, 41 HASTINGS L.J. 777, 781–91 (1990).
99 It is worth noting that there is a long tradition in the United States of using the courts as forums for protest. Lobel, supra note 79, at 493–510.
100 Cain, supra note 93, at 1 (“Taxpayers generally do not . . . use public interest law firms in the ways that civil rights groups do.”); but see supra note 83.
with arguments tailored to their individual situations.\textsuperscript{103}

This diffuse, individualized approach to tax controversies is perfectly suited for what has been called the “war of the flea”:

“Analogically, the guerrilla fights the war of the flea, and his military enemy suffers the dog’s disadvantages: too much to defend; too small, ubiquitous, and agile an enemy to come to grips with. If the war continues long enough—this is the theory—the dog succumbs to exhaustion and anemia without even having found anything on which to close his jaws or to rake with his claws.”\textsuperscript{104}

While it is easy enough for the government to close its jaws on a single lawsuit,\textsuperscript{105} imagine what would happen if thousands of domestic partners in California\textsuperscript{106} were each to file federal income tax returns splitting their earned income in accordance with the

\textsuperscript{103} See Cain, supra note 93, at 1 (“Taxpayers generally do not file class actions [or] share the expense of litigation . . . .”).

While it is possible to prosecute a class action lawsuit seeking a refund of taxes, refund claims have been described as “individualized” and “particularly ill-suited for class certification” because of the need, under I.R.C. § 7422(a) (2006), for each class member to pay the amount of her tax deficiency and then timely file a claim for a refund. Saunooke v. United States, 8 Cl. Ct. 327, 330 (Cl. Ct. 1985); see also Appoloni v. United States, 219 F.R.D. 116 (W.D. Mich. 2003) (narrowing the definition of the “class” in a class action lawsuit for refund of Federal Insurance Contribution Act taxes so that it would meet the requirements of I.R.C. § 7422(a)); Rose v. American Airlines, Inc., 331 F. Supp. 77, 79 (N.D. Ill. 1971) (class action claim in lawsuit challenging airport taxes stricken for failure to meet the I.R.C. § 7422 requirement of individual refund claims filed by each member of the class); Agron v. Ill. Bell Tel. Co., 325 F. Supp. 487, 488 (N.D. Ill. 1970) (same with respect to class action claim in lawsuit challenging telephone taxes); McConnell v. United States, 295 F. Supp. 605, 606 (E.D. Tenn. 1969) (refusal to certify class in income tax refund action).

Subject to very narrow exceptions, the rules of the U.S. Tax Court do not appear to contemplate the possibility of a class action suit challenging the Service’s assertion of a tax deficiency. Compare Tax Ct. R. 60(a)(1) (“A case shall be brought by and in the name of the person against whom the Commissioner determined the deficiency (in the case of a notice of deficiency) or liability (in the case of a notice of liability) . . . .”) and id. R. 61(a) (“No person, to whom a notice of deficiency or notice of liability has been issued, may join with any other such person in filing a petition in the Court, except as may be permitted by Rule 34(a)(1).”) and id. R. 34(a)(1) (permitting a joint petition to be filed only by a husband and wife or by multiple persons who are sent a single notice of deficiency or liability by the Service) with Tax Ct. R. 215 (allowing joinder of parties in declaratory judgment actions concerning retirement plans and estate tax installment payment actions), and id. R. 226(a) (allowing joint petitions in disclosure actions).

\textsuperscript{104} Fraser, supra note 98, at 782 n.20 (quoting ROBERT TABER, THE WAR OF THE FLEA: A STUDY OF GUERRILLA WARFARE THEORY AND PRACTICE 29 (1970)).

\textsuperscript{105} See Infanti, Tax Protest, supra note 87 (describing the government’s treatment of a gay man who protested the treatment of lesbian and gay couples under the federal tax laws).

\textsuperscript{106} As of February 28, 2006, there are 37,283 registered domestic partnerships in the State of California. E-mail from dp@ss.ca.gov to Anthony C. Infanti, Associate Prof. of Law, Univ. of Pittsburgh Sch. of Law (Feb. 28, 2006) (on file with author) (this e-mail was sent in response to a request for information that I made using an online form on the domestic partner registry web page of the California Secretary of State). The State of California does not provide a breakdown between same-sex and different-sex domestic partnerships, as that “information is not required by law and cannot be asked for on the Declaration.” Id. As a result, some of these partnerships may be between different-sex couples, provided that at least one member of the couple is over the age of 62. CAL. FAM. CODE § 297(b)(5)(B) (Deering 2006).
U.S. Supreme Court’s decision in *Poe v. Seaborn* and were openly to invite the Service to audit their returns and to challenge their interpretation of the law. Then consider what would happen if, at the same time, thousands of married same-sex couples in Massachusetts were to file joint federal income tax returns, asserting that DOMA is unconstitutional, and likewise were openly to invite the Service to audit their returns.


108 To be clear, I contemplate here the filing of returns that on their face challenge the current application of the tax laws to same-sex couples. Such a transparent challenge could be accomplished by filing the appropriate disclosure forms, see infra note 113, and including a cover letter with the return that explains the precise nature of the challenge. See, e.g., Drucker v. Commissioner, 697 F.2d 46, 47–48 (2d Cir. 1982) (explaining the steps taken by taxpayers who wished to openly challenge the constitutionality of the marriage penalty). Surreptitious challenges simply would not have the same effect: Although the Service would likely uncover the income-splitting relatively quickly and easily because of the mismatch between the amount of earned income reported on the tax return and the amount of earned income reported to the Service by the taxpayer’s employer, the Service would be far less likely to detect the filing of a joint federal income tax return by a married same-sex couple from Massachusetts, as suggested in the text below. See Warren Rojas, *Financiers: Bush Marriage Crusade May Only Bolster Tax Inequities*, 102 TAX NOTES 1597, 1598 (2004) (indicating that it would be difficult for the Service to detect joint filing by same-sex couples). In either case, open challenges are necessary to bring about the desired adverse impact on the Service, because achieving that impact depends on the Service’s immediate awareness of the challenges.

109 See Cain, supra note 93, at 3, 16 (indicating that the time is ripe for such a challenge); see also Schwanhausser, supra note 8 (prior to the issuance of the memorandum described in the next paragraph, indicating the level of frustration among taxpayers and tax preparers as they attempt to grapple with this issue as well as the difference of opinion among tax experts over whether income-splitting would be accepted by the courts); Mark Schwanhausser, *Income-Splitting on Taxes A Gray Area*, SAN JOSE MERCURY NEWS, Feb. 23, 2006, at 15A (same).

On February 24, 2006, the Service issued Chief Counsel Memorandum 20060838, available at http://www.irs.gov/pub/irs-wd/0608038.pdf, which addresses the reporting of earned income by California domestic partners. The Chief Counsel’s Office opined that domestic partners must each report their earned income separately, taking the position that *Poe v. Seaborn* applies only to married couples. For criticism of this memorandum, see Cain, supra note 93, at 12–15, and Dennis J. Ventry, Jr., *No Income Splitting for Domestic Partners: How the IRS Erred*, 110 TAX NOTES 1221 (2006). For a history of the attempts to get the Service to speak on this issue, see Cain, supra note 93, at 2–3, and Ventry, supra, at 1221 n.2.

This opinion was not only provided late—well into tax season and only a matter of weeks before the April 15 deadline for filing returns—but also came in a form that is prohibited by law from being cited as precedent. I.R.C. § 6110(k)(3) (2006); see also I.R.S. Chief Couns. Mem. 20060838 (Feb. 24, 2006) (“In accordance with § 6110(k)(3) this advice may not be used or cited as precedent.”). If anything, this advice (i.e., advice that is not advice—to borrow the strikethrough that is sometimes used by Agamben, see, e.g., AGAMBEN, EXCEPTION, supra note 19, at 32–40, and Derrida, see, e.g., DERRIDA, supra note 25, at 19, 44, but which did not originate with them, see Gayatri Chakravorty Spivak, *Translator’s Preface* to DERRIDA, supra note 25, at ix, xiv—the obvious choice of form for advice to be given to those consigned to tax “limbo”) only heightens the need for, and the potential success of, the guerrilla warfare tactics described in the text above. See Mark Schwanhausser, *Domestic Partners to Be Taxed as Singles*, SAN JOSE MERCURY NEWS, Feb. 28, 2006, at 1C (“The memo from the office of chief counsel doesn’t have the force of law, but tax experts say it sends a signal that the IRS intends to challenge domestic partners who split their income based on rules for married couples who file separately.”).

110 The most recent figures are for the period from May 17, 2004 (when same-sex marriage first became legal in Massachusetts) through December 31, 2004. During that period, 5,994 same-sex couples were married in Massachusetts: there were 2,123 male-to-male marriages and 3,871 female-to-female marriages. E-mail from Kevin Foster, Mass. Dep’t of Pub. Health, to Anthony C. Infanti, Assoc. Professor of Law, Univ. of Pittsburgh Sch. of Law (Feb. 21, 2006) (on file with author).

and to challenge their interpretation of the law.\textsuperscript{112}

The Service should be accustomed to, and adequately equipped for, individual challenges to its interpretation of the tax laws; in fact, it has promulgated specific regulations concerning, along with the necessary forms for reporting, tax return positions that challenge rules or regulations.\textsuperscript{113} However, the nearly simultaneous arrival of thousands upon thousands of these challenges would, I imagine, leave the Service nonplussed (to say the very least).\textsuperscript{114} We could expect that this initial strike against the law would cause a disruption similar to that caused by a computer virus that is designed to bombard a single website with information until it becomes so overloaded that it is forced to be shut down temporarily.\textsuperscript{115} After this initial disruption, the Service would be saddled with thousands of individual audits and the resulting court cases, which could drag on for years and significantly drain government resources.

With a bit more effort, we could turn the law even more fully against itself, while at the same time facilitating the participation of as broad a swath of the lesbian and gay community as possible in this opening salvo of the guerrilla war. We could encourage

\begin{quote}
Even though a state may recognize a union of two people of the same sex as a legal marriage for the purposes within that state’s authority, that recognition has no effect for purposes of federal law. A taxpayer in such a relationship may not claim the status of a married person on the federal income tax return.
\end{quote}

Clause and Due Process Clause prohibit Congress from enacting DOMA); Mark Strasser, \textit{Ex Post Facto Laws, Bills of Attainder, and the Definition of Punishment: On DOMA, the Hawaii Amendment, and Federal Constitutional Constraints}, 48 SYRACUSE L. REV. 227 (1998) (arguing that DOMA violates the Bill of Attainder Clause); Mark Strasser, \textit{Loving the Romer Out for Baehr: On Acts in Defense of Marriage and the Constitution}, 58 U. PITT. L. REV. 279 (1997) (arguing that enactment of DOMA exceeds Congress’ power under the Full Faith and Credit Clause, violates the right to interstate travel, and does not meet the relevant standard for displacing state domestic relations law).\textsuperscript{112} In response to a letter from a conservative, “pro-family” organization urging it to investigate and prosecute any same-sex couples that might attempt to file joint federal income tax returns, the Service reaffirmed the application of DOMA to the tax laws and summarized its position on joint filing by same-sex couples as follows:


\begin{quote}
What is essential is that, every time refugees represent not individual cases but—as happens more and more often today—a mass phenomenon, both these organizations and individual states prove themselves, despite their solemn invocations of the “sacred and inalienable” rights of man, absolutely incapable of resolving the problem and even of confronting it adequately.
\end{quote}

low- and middle-income as well as elderly lesbians and gay men to avail themselves of volunteer income tax return preparers who are trained and supported by the Service when preparing their returns challenging the Service’s interpretation of the tax laws. The Service sponsors two volunteer tax return preparation programs: for low- and middle-income taxpayers, the Volunteer Income Tax Assistance program provides help in preparing basic tax returns; for the elderly, the Tax Counseling for the Elderly program provides free tax counseling and help in preparing basic tax returns. The sites where help is available are located throughout the country, and the Service will provide taxpayers with the location of the closest site when they call its toll-free telephone number. Following the initial disruption, the extended audit and litigation process would likely require the participating lesbian and gay taxpayers to employ accountants and/or attorneys to work on their behalf—all of whom would obviously be instructed to be as unaccommodating to the Service as possible (e.g., by refusing to extend the statute of limitations to allow more time for audit). For those with means, this would entail a financial sacrifice. Those who would find this sacrifice unduly burdensome could turn for help to attorneys and accountants who are willing to work on a pro bono basis. Again, however, an opportunity arises to turn the law even more fully against itself (and, at the same time, to facilitate the participation of as broad a swath of the lesbian and gay community as possible): Low-income taxpayers could contact one of the many independent legal clinics that provide professional assistance in disputes with the Service. These clinics receive financial support from the Service and can be found in every state and the District of Columbia.

If this initial strike were to prove successful, lesbians and gay men could then begin to educate the masses concerning their tax grievances. These educational activities could be followed by the expansion of our guerrilla activities into other areas of the law. Naturally, we would need to adapt our tactics as we branch out to other areas of

---


117 Id.

118 No one ever said that subversion of the privileging of heterosexuality in our society would be cheap or easy:

> These figures push the aporia of sovereignty to the limit but still do not completely free themselves from its ban. They show that the dissolution of the ban, like the cutting of the Gordian knot, resembles less the solution of a logical or mathematical problem than the solution of an enigma. Here the metaphysical aporia shows its political nature.


119 Many tax lawyers ache to do pro bono work, but are hard put to find cases that allow them to put their tax skills to use. A good place to start looking might be the listing of pro bono programs maintained by the American Bar Association on its website. Am. Bar Ass’n, Directory of Pro Bono Programs, [link](http://www.abanet.org/legalservices/probono/directory.html#). To obtain information about accountants who provide pro bono services, see Accountants for Pub. Interest, [link](http://www.geocities.com/api_woods/api/apihome.html).


121 Fraser, *supra* note 98, at 786 (“After [a guerrilla band’s] initial successes, it sets about to educate the masses . . . .”).
the law, both because other areas might require different approaches and because the
government would probably have adapted to (and, therefore, would likely be
anticipating) the tactics that we used in the tax context. In each case, however, the goal
would be the same—to provoke the closure of yet another already open door to the law.

THOUGHT-PROVOKING, PROVOKING THOUGHT

For decades, we have pursued a legal strategy that engages the law on its own
terms and that has afforded us only the most tentative of progress toward the unqualified
acceptance that we seek from the heterosexual majority. To justify our efforts, we have
reassured ourselves that, over time, this strategy will win out. But what if all this time we
have merely been patiently waiting outside the already open door to the law, allowing
ourselves to be cast as homines sacri and to be placed in the dangerous position of
serving as the exception that gives meaning to the general rule of heterosexual privilege?
What if the only way out of this dangerous position is to provoke the closure of the door
to the law and the concomitant conversion of our virtual state of exception into a real
one? What if turning the law against itself—using the law as a tool in guerrilla warfare
against those in the heterosexual majority who wield power—is the only effective means
of clearing the way for a meaningful and thorough reconsideration of the appropriate
relationship between sexual orientation and legal and social norms? While these
questions may make us feel nervous and uncomfortable (as any truly thought-provoking
questions should), we must open ourselves to thinking the possibility that the time has
arrived for lesbians and gay men to radically alter their approach to effecting legal and
social change.
THE PRICE OF CONFLICT:
WAR TAXATION IN AMERICAN HISTORY

The Civil War and Reconstruction

Steven A. Bank*

I. Introduction

The Civil War provides perhaps the quintessential example of wartime taxation. First, the war created immediate need. As a result of the secession of the southern states and the outbreak of hostilities, both the Union and the newly-constituted Confederacy had to raise funds to erect or greatly expand their military forces and to develop the administrative apparatus to conduct the war. For the Union, this need followed an almost unheard of four consecutive years of deficits, while, in the case of the Confederacy, the need was on top of the expenses associated with creating a government from scratch. Second, the war disrupted traditional sources of governmental support. Tariff income, which had constituted more than ninety percent of federal revenues during the 1850s,1 was less reliable both because of naval blockades and the economic dislocation occasioned by the war. Debt financing was available, but costly due to a precipitous drop in the Union credit rating that accompanied the growth in the deficit. In the Confederacy, the situation was even worse because of the lack of established access to capital or specie to support the issuance of paper money. Finally, the war exacted immediate and visible physical and material sacrifices. These included both the injuries and loss of life from the

* Professor, UCLA School of Law. Thanks to David Martin for his research assistance.
battlefield as well as the government confiscation of property and the destruction of the American landscape that took place in many areas. These sacrifices imposed a sort of implicit taxation that helped drive the direction and shape of the explicit tax system as a whole.

The confluence of these three factors—significant financial need, reduced access to traditional sources of revenue, and the imposition of great sacrifices outside the explicit tax system—led to large-scale experimentation and innovation in the field of public finance. As Joseph Hill remarked, some of the measures adopted were ones “which the boldest innovator would hardly have dared propose in times of peace; but under the stress of war they were enacted, sometimes with little opposition and little discussion.”\(^2\) Virtually all goods were subject to some form of internal excise tax, stamp tax, or other levy. Perhaps most notably, the Union collected an income tax for the first time in 1862. Not only was this tax based upon the ability to pay rather than the necessity to consume, it was adopted with a progressive graduated marginal rate structure. Supplementing the progressivity of this income tax was the enactment of the nation's first graduated rate inheritance or succession tax.\(^3\) The Civil War and Reconstruction also introduced modern notions of third-party collection techniques, including payroll withholding for salaried government employees and the withholding of taxes from dividends and interest paid by businesses in industries such as transportation, finance, and insurance. Finally, there were a number of novel measures that mitigated the impact of the new income tax, such as a deduction for rental payments that included a

---


\(^3\) A graduated stamp tax on legal transactions such as the receipt of legacies and the probate of wills was imposed between 1797 and 1802, but there had not been a tax on inheritance per se. Randolph Paul, \textit{Taxation in the United States} (Boston: Little, Brown & Co., 1954): 15-16.
deduction for the imputed rent paid by home owners and a reduced rate on the interest paid from Treasury securities, which was designed to spur the purchase of war bonds.

This is not to say that there was unanimous agreement that great fiscal sacrifice was necessary or that human and material sacrifice was a necessary precondition to tax innovation and reform. The Confederacy, prompted by optimism about the duration of the war and fear of centralized power, initially held off on employing some of the new methods of finance. To many in the South, a weak national government was part of what they sought to preserve. This was particularly true in the case of taxation given its potential to destroy the economic basis for slavery. As a consequence, the Confederate government relied on printing money through issuances of unsupported Treasury notes, with disastrous consequences. By contrast, after an initial adjustment period of its own, the Union more quickly embraced new avenues of taxation, including a modestly progressive income tax to balance out the regressivity of the previously-dominant tariff tax. It ramped up internal taxation of goods and products, creating the administrative apparatus that served as the forerunner of the modern Internal Revenue Service. While the Confederacy eventually embraced similar measures later in the war, it was too late to change its fortunes.

Moreover, even the Union appeared to abandon the concept of fiscal sacrifice when the crisis began to diminish. While the Confederate tax program obviously ended abruptly with the conclusion of the war, the Union income tax and many of the internal taxes were also allowed to expire after Reconstruction. Relatively high customs duties continued following the demise of the income tax, but this was only after a planned reduction of the rates was repealed in the aftermath of the Panic of 1873. Furthermore, it
was the same basic tariff structure that had been in place before the war and it continued to fund high percentages of federal revenues as it had in the 1850s.

Notwithstanding this retrenchment, the Civil War tax program still had a notable legacy. First, some of the new internal excise taxes continued under the auspices of the now permanent Office of Internal Revenue — most notably "sin" taxes on alcohol and tobacco. Not only did this administrative apparatus for administering internal taxation survive the war and continue to this day, but these internal taxes at least partially helped to finance reconstruction efforts, public works projects, and pension and disability benefits for Union veteran soldiers, the latter of which was effectively one of the first large-scale progressive social welfare programs. Just as significant, however, is that the Civil War tax experience fundamentally changed both the nature and tenor of the debate concerning public finance. While direct taxes such as the income tax had gradually become a part of the debate by the end of the 1850s, the Panic of 1857 ended any substantial movement in this regard. After the Civil War, with the divergent financial results of the Union and Confederacy, the income tax became a fixture in the legislative debate throughout the 1870s and 1880s. Between 1873 and 1879 alone, fourteen different income tax bills were introduced in Congress and it continued to have a central role during the 1880s and early 1890s. Significant tariff reform even became a more realistic possibility in light of the Civil War evidence that there were means for reducing reliance on customs duties. Ultimately, the experience paved the way for the subsequent

---

adoption of an income tax, briefly in 1894 and more permanently in 1913. Thus, while the apparent end to the fiscal crisis undercut the demand for individual sacrifice in taxation, the war had an enduring effect on the American tax system.

II. Antebellum period

During the first half of the nineteenth century, the predominant source of federal revenues came from the tariff. While not technically a tax, this method of finance nevertheless imposed an indirect burden on domestic consumers. This is because the customs duties levied on foreign producers who sought to export their goods to this country were passed along to consumers in the form of higher prices. Moreover, the higher prices often were not limited to the goods subject to such duties. In cases where the foreign goods competed with domestic products, the domestic manufacturers also increased the prices on their goods, albeit to a level just below that of the foreign products. Typically, this was by design as it protected the domestic industries from foreign competition while allowing them to reap higher profits from the sale of their goods.

Notwithstanding the well-established presence of tariffs during the period, there was substantial disagreement over their use. This disagreement generally, although not exclusively, followed sectional boundaries. In the North, where domestic manufacturing was prevalent, the preference was for protectionist tariffs that would shield American producers from foreign competition. By contrast, in the South, where rural planters were net exporters to Europe, the preference was for a primarily free trade-based approach. For these planters, protectionist tariff policies would only serve to increase the price of
goods they purchased from northern manufacturers without increasing the price they could obtain for their agricultural products sold abroad. Under this free trade view, tariffs could be used to raise revenue, but only if they were limited to items not produced or available in America, which therefore did not have a protectionist effect. This sectional divide had long-standing roots. South Carolina went so far as to threaten to secede over the "Black Tariff," or "Act of Abominations," enacted in 1828, which imposed rates of almost 50 percent on some goods.6

Gradually, the pressure for free trade began to win out. The backlash to the heavily protectionist Tariff of 1828 spread beyond South Carolina and soon gave way to lower duties. The accumulation of a government surplus helped spur the enactment of the Tariff of 1857, which dramatically reduced protective tariffs to their lowest level in forty years.7 Duties were cut across the board by twenty to twenty-five percent, with the highest rate set at twenty-four percent.8 While the protest from formerly protected industries crossed over sectional boundaries,9 the tide had appeared to turn toward a new era of government finance.

8 Id.
9 For instance, the reduction in sugar duties caused much consternation in the South. The New Orleans Bulletin protested that "protection has been extended to the iron, cotton, and woolen interests of the North - - why withhold it from the sugar interest of the South?" Quoted in "The Sugar Tax Again," National Era, Vol. XI, No. 526, Jan. 29, 1857, at 18 (APS Online) (also reporting "the sectional clamor got up by a few Southern papers against the repeal of the tax."). The National Era, an abolitionist weekly that reported broadly on economic events, responded that it was simply not a suitable industry for protection: "[a]ll the protection in the world cannot bring [sugar production] up to the demand -- cannot enable the planters to render us independent of foreign supplies. We have always been, and shall always be, dependent upon other countries for the larger portion of our consumption. . . . It is simply in the nature of an onerous and a perpetual bounty." "The Sugar Tax Again," National Era, Vol. XI, No. 526, Jan. 29, 1857, at 18 (APS Online). Perhaps to bolster Southern support for protectionism, some Northern newspapers joined the cause in favor of tariffs on sugar, reportedly leading the House Ways and Means Committee to propose reducing the sugar tax from thirty to ten percent instead of repealing it outright. "Sugar Tax," National Era, Feb. 5, 1857, Vol. XI, No. 527, Feb. 5, 1857, at 22 (APS Online).
Given the reduction in protective tariff rates, some observers began to turn their attention to the possibility of using direct taxation as a vehicle for phasing out revenue tariffs as well. *DeBow's Review*, a southern journal focused on agricultural commerce, ran a series of hypothetical dialogues between a farmer, planter, and a politician on the benefits of free trade and direct taxation.\(^{10}\) After the first dialogue, which discussed the evils of the tariff, the parties turned to the question of what could replace the tariff as a source of revenue. The Planter declared "I am willing to pay a tax to support the Government according to my ability, because I believe it to be fair, just and equal." The Farmer did not immediately disagree, but reported that the Politician "says if we abolish the tariff and establish free trade and direct taxes, the Government will send swarms of tax gatherers to vex, harass, and eat us out; and it will cost a much higher per cent. to collect direct taxes than to collect duties on imports by the tariff." This, responded the Planter, would only cause the Politician to be voted out of office, thereby leading to the reduction of rate and moderation of intrusiveness. The experience with property taxation at the state level was cited as evidence for this point. The tariff, explained the Planter, is much worse because "no man knows how much he pays and very few have any idea how much he is taxed, and many do not know they are taxed at all. Consequently, nobody watches the Government, and the public funds are squandered recklessly in the wildest schemes of extravagance, or worse still, applied to the basest purposes of corruption to bribe your members to betray you interest, or to reward the favorites of the party in power."

Unfortunately for its supporters, this potential shift toward direct taxation was abandoned as a result of economic dislocation caused by the Panic of 1857. After a dozen years of economic growth, fueled in large part by speculative investment, the house of cards came tumbling down towards the end of 1857. While there have been many explanations offered for the origins of the Panic, its impact was fairly evident. A decline in securities prices and land values in the West precipitated the demise of factories, businesses, and construction. Hundreds of thousands of workers lost their jobs in the process. Furthermore, the depression had a significant effect on government revenues. Imports fell by over twenty-five percent in a one year period, leading customs revenues to drop by one-third.

Between 1856 and 1861, revenues dropped by almost $25 million while spending only dropped by less than $2 million. For the first time in a decade, the federal government ran a deficit. Even though the depression itself was relatively short, with much of the economy returning to normal by early 1859, the federal deficit continued to mount.

There was nothing apparent at the time or even now to suggest that this chain reaction of events was rooted in or exacerbated by the low tariffs enacted in 1857, but

---

14 Ashley, supra note 6, at 177-78; Frederic C. Howe, *Taxation and Taxes in the United States under the Internal Revenue System, 1791-1895* (New York: Thomas Y. Crowell, 1896): 51; Smith, supra note 7, at 3.
16 Prior to 1858, the last time the government ran a deficit was in 1849. U.S. Dept. of Commerce, Historical Statistics of the United States, 1789-1945 (1949): 297, Series P 89-98.
17 Id. at 191.
18 Ashley, supra note 6, at 178-79.
19 Id. at 177.
that did not stop protectionists from seizing upon the economic crisis to make their case. Horace Greeley wrote in the New York Tribune that "[n]o truth of mathematics is more clearly demonstrable than that the ruin about us is fundamentally attributable to the destruction of the Protective Tariff."\textsuperscript{20} Such rhetoric was used to rally labor against the reduced tariffs. In Pennsylvania, where there was concern that the lower duties adopted on iron would subject the iron industry to stiff competition from British firms, Republicans promised to laborers that they "shall be protected against the pauper labor of Europe [noting that a higher tariff would] "give employment to thousands of mechanics, artisans, laborers, who have languished for months in unwilling idleness."\textsuperscript{21} In the 1860 Presidential campaign, the Republican platform included a plank promising an "adjustment" of tariff rates "to encourage the development of the industrial interests of the whole country" and "secure to the workingman liberal wages."\textsuperscript{22}

In light of the rising federal deficits and the pressure to resume protection in aid of the economy, Congress in 1860 considered a bill to restore customs duties to their pre-1857 level.\textsuperscript{23} This bill faced significant obstacles, however, due to the Democratic majority in the Senate. If the Southern states had not seceded at the end of the year, the bill may not have passed at all, forcing Congress to resort to some other mechanism for closing its deficit. While direct taxation may not have been on the table, there was still interest in continuing the reform of the revenue and tariff system that had taken place over much of the previous two decades, culminating in the passage of the Tariff of 1857.

\textsuperscript{21} Lebanon (Pa.) Courier, quoted in McPherson, supra note 13, at 192.
\textsuperscript{22} Quoted in McPherson, supra note 13, at 221.
\textsuperscript{23} Ashley, supra note 6, at 178-79. One additional factor prompting the turn to increased tariffs was the fact that the impending political crisis had disrupted the market for government bonds. Patterson, supra note 15, at 38.
This interest subsided, though, with secession. The withdrawal of several Democratic senators from Southern states gave Republicans the majority control necessary to enact the new higher tariff rates, which came into effect in March of 1861.24 This pro-tariff victory would be short-lived, however, as the outbreak of the Civil War provided the impetus to revisit the direct tax debate on both sides of the Mason-Dixon line.

III. Civil War

At the outset of the Civil War, both the North and South faced severe financial difficulties. The North was still reeling from the Panic of 1857. Four straight years of budgetary deficits had left the newly-elected President Abraham Lincoln with an aggregate debt of approximately $75 million.25 The South, of course, had to create a government and raise an army, all without the benefit of established credit or currency.26 Despite these similarities, during a crucial period in the war the two governments chose to finance their activities very differently. Once it became apparent that the hostilities would not be short, the Union resorted to directed taxation and internal excises to supplement its bonds and treasury notes. The Confederacy, by contrast, avoided such taxing measures for much longer. This difference in their approach to taxation and

24 Id. at 179.
26 According to one account, the Confederate access to capital was so limited that it initially could only obtain credit for operating expenses when its Secretary of the Treasury, Christopher Memminger provided a personal guarantee to a local bank. Paul, supra note 3, at 17.
finance may have reflected the divide between the North and South over their respective attitudes toward, and their willingness to sacrifice for, the creation of centralized power.27

A. Confederacy

In the new Confederate states, there was initially optimism that the war would end quickly.28 One Southern journal predicted that "[t]he strange infatuation by which the people of the North are deluded to believe that it is in their power to subjugate us and force us back into political union with them and under their Government, cannot endure very long. A few weeks, or at most a few months more, with the experience they have already had, can scarcely fail to bring them to a sober perception of the insuperable difficulties of an enterprise which has appeared to them to be so easy to be accomplished."29 Thus, the journal advised that "it is possible, and perhaps even not altogether improbable" that any tax enacted "may never be collected."30

29 "The War Tax," DeBow's Review and Industrial Resources, Statistics, etc. Devoted to Commerce VI(4) (Oct/Nov 1861), 436. Several years later, President Jefferson Davis confirmed that this overconfident attitude about the duration of the war adversely affected the shape of revenue policy:

At the commencement of the war we were far from anticipating the magnitude and duration of the struggle in which we were engaged. The most sagacious foresight could not have predicted that the passions of the Northern people would lead them blindly to the sacrifice of life, treasure, and liberty in so vain a hope as that of subjugating thirteen independent states inhabited by millions of people, whose birthright of freedom is dearer to them than life. A long exemption from direct taxation by the General Government had created an aversion to its raising revenue by any other means than by duties on imports, and it was supposed that these duties would be ample for current peace expenditure, while the means for conducting the war could be raised almost exclusively by the use of the public credit.

3 Journals of the Confederate Congress 442 (Dec. 8, 1863).
30 Id.
This attitude, plus the decline of tariff revenues due to the Union blockade of Confederate ports in the spring of 1861, may have helped propel an early strategy that relied almost exclusively on loans and notes. According to one study of Confederate finances, bonds and treasury notes accounted for ninety-nine percent of total revenues between July 20 and November 16, 1861. This dipped only slightly to ninety-six percent for the 1862 fiscal year and was still as high as ninety-five percent for the first nine months of 1863.

The lone attempt to impose a direct tax during the first two years of the war was less than successful. After the first Battle of Bull Run in July of 1861, the Provisional Congress adopted a direct tax of one-half of one percent on all property, including real estate, slaves, and personal property. This was necessary to support the borrowing of funds, since the Confederacy otherwise lacked adequate access to capital. Rather than impose the tax on the people directly, though, which would have required the erection of its own collection apparatus, the Confederacy generally permitted individual state governments to assume the obligation and pay it themselves. As a consequence, few of

---

31 Paul, supra note 3, at 18; "A War Tax is a Necessary Measure," DeBow's Review and Industrial Resources, Statistics, etc. Devoted to Commerce VI(4) (Oct/Nov 1861): 436 ("It is obvious that during the continuance of the war with the United States, and the blockade of our ports, our foreign commerce must be interrupted and little or no revenue can be derived from duties on imports."). Because of the blockade, customs duties were a mere one percent of total revenues in the year ending February 1862 and declined rapidly from there. J.C. Schwab, "The Finances of the Confederate States," Political Science Quarterly, 7(1) (March 1892), 38, 40. The $3.5 million raised over the course of the war fell quite short of Confederate projections of $25 million in the first year of operation. See Ball, supra note 28, at 204-07.


33 Id.


35 "The Finances of the Confederate States," N.Y. Times, Mar. 16, 1861, at 4 ("The weakness of the Confederate States in their finances, is shown by the fact that the proceeds of a special tax is mortgaged for the payment of their first loan. . . . If these States commence by mortgaging the best source of revenue for their first loan, what is to guarantee the second, that must speedily follow?"). One of the reasons the South lacked the ability to borrow is that some of the states had actually repudiated their own debts already. Id.

36 See Ball, supra note 28, at 226 ("To mobilize a staff of collectors and assessors capable of appraising and levying direct taxes, to train them in their arduous duties, and to devise the appropriate instructions and regulations required far more will, and, to a lesser degree, more skill than the Confederacy possessed.").
the states actually collected any tax at all, choosing instead to issue their own treasury notes or to borrow from their local banks, thus serving only to further increase the aggregate Confederate debt burden and reduce the value of state and federal treasury issuances. During its two years of operation, only $20 million was raised and little of it came from the imposition of the Confederate tax itself.

By relying upon the issuance of unsupported notes and the federal and state level, the Confederacy essentially printed money to finance the early stages of the war. This had disastrous effects on the Southern economy, causing rampant inflation. The excessive issuances of Treasury notes and bonds had led to a level of hyperinflation unseen since the Revolutionary War, with the price of gold in Confederate dollars rising to as high as sixty-one dollars. There was a corresponding effect on the price of commodities. Prices of such staples as wheat, bacon, and flour rose by as much as

---

37 Only South Carolina actually collected the direct tax itself in cash, although Mississippi collected taxes to make interest payments on bonds issued to fulfill their quota. Texas confiscated property located in the state, but owned by Northerners, Alabama borrowed from state banks, and all other states issued securities to obtain the funds. Eugene M. Lerner, “The Monetary and Fiscal Programs of the Confederate Government, 1861-65,” *Journal of Political Economy* 61(6) (December 1954): 506, 509.

38 Seligman, *supra* note 34, at 482-83.

39 To make the transition to taxation more palatable, the tax as passed was only supposed to be collected starting in May of 1862, almost a year after the law was enacted. See Razaghian, *supra* note 27, at 13.


42 This inflation financing had damaging effects in the South for years afterward. See Peter Temin, "The Post-Bellum Recovery of the South and the Cost of the Civil War," *Journal of Economic History*, Vol. 36, No. 4: 898-907.

43 Id. at 22. See James L. Sellers, "An Interpretation of Civil War Finance," *American Historical Review* 30(2) (January 1925): 282 ("December 1863, in the very middle of a four-year struggle, the value of the Confederate dollar had fallen so low that it was worth only 5 cents in gold."); John Christopher Schwab, "Prices in the Confederate States, 1861-65," *Political Science Quarterly* 14(2) (June 1899): 281, 286 ("The relatively rapid increase in the issue of notes after August, 1862, during the last months of August 1863 and again during the last months of the war, correspondingly drove up the [gold] premium at those three different times.").

44 Schwab, *supra* note 43, at 293 ("Each unusual rise in the gold premium was not coincident with, but was followed by, a general rise in the prices of commodities.").
2,800 percent between 1863 and 1865.\textsuperscript{45} Local differences in these prices of commodities encouraged rampant speculation, which only fanned the flames of economic distress.\textsuperscript{46}

The failure of this direct tax appeared to reflect a more deep-seated distrust of centralized taxation that may further explain Confederate financing decisions. While some have concluded that Confederate economic and social policies suggest an embrace of a centralized post-secession government,\textsuperscript{47} Confederate tax policies, at least during the crucial first half of the war, pointed in the opposite direction. Edwin Seligman, a Columbia economics professor and noted nineteenth century expert on taxation, observed that "the aversion of the people to direct taxation was so great as to have rendered the [Confederate] war tax practically nugatory."\textsuperscript{48} According to Rose Razaghian's analysis of roll call votes in the Provisional Congress, this reflected the antebellum concerns about centralized power. "Congressmen who were the most deeply committed to protecting states' rights and slavery were more likely to rely on financial policies that limited the aggrandizement of federal power, issuing loans and notes, and were more likely to oppose policies that empowered the federal government, in particular taxes."\textsuperscript{49} The view appeared to be that the Confederacy could proceed without requesting deep fiscal sacrifices from its citizens.

\textsuperscript{45} R. Neil Fulghum, "Moneys for the Southern Cause," Documenting the American South: The Southern Homefront, 1861-1865, available at http://docsouth.unc.edu/imls/currency/index.html (last visited 10/6/05). In some cases, the rise in these prices was due as much or more to scarcity than to the issuance of paper notes. Current, supra note 45, at 23.
\textsuperscript{46} Schwab, supra note 43, at 294.
\textsuperscript{48} Id. at 483.
\textsuperscript{49} Razaghian, supra note 27, at 4.
The notion was that debt financing required “minimal state apparatus and could be administered by few bureaucrats with limited resources.”\textsuperscript{50} Taxation, by contrast, not only directly involved the central government in the lives of its citizens through the expansion of a federal bureaucracy created to administer the tax, but it was more likely to become entrenched.\textsuperscript{51} Alabama’s governor, A.B. Moore, epitomized this attitude in a letter to the Confederate Congress explaining why the states should not cooperate in collecting the Confederacy’s direct tax:

\begin{quote}
The State should never concede to the General Government the exercise of powers not delegated to the Constitution, and they should never, except in cases of absolute necessity, consent to exercise powers or to perform duties which do not properly belong to them. As a general rule it is dangerous in its tendencies, and precedents of this character are to be avoided. The collection of this tax by the State would be an onerous and unpleasant duty, as it imposes upon the State the necessity of enforcing the laws of the Confederate Government against her own citizens . . .\textsuperscript{52}
\end{quote}

For a people "who had even more than the typical American's resistance to taxation," the hesitance to embrace taxation immediately was understandable.\textsuperscript{53}

A contributing factor in the Confederate reluctance to impose taxation was the inexperience of Confederate Treasury officials. President Davis’ Secretary of the Treasury, Christopher G. Memminger, was only marginally acquainted with contemporary economic thought and lacked the forceful personality to push through tax measures even when he deemed them necessary.\textsuperscript{54} This does not mean he failed to

\begin{flushleft}
\textsuperscript{50} Id. at 7.
\textsuperscript{51} Id. at 8-9.
\textsuperscript{54} See Ball, \textit{supra} note 28, at 9. But see Razaghian, \textit{supra} note 27, at 22 (noting that several high ranking Treasury officials joined the Confederate Treasury, although often in more administrative than policy capacities).
\end{flushleft}
advocate in favor of relying on increased taxation. In February 1862, he pleaded before the Confederate Congress that “[t]he war tax has already put in motion all the machinery requisite for levying a tax. It has selected those articles which can best bear the burden, and it levies on their value the very moderate rate of one-half of one per cent. The simplest of all plans, therefore, would be an increase of this tax.” Memminger’s counsel was ignored at the time, however, which is perhaps not surprising given President Davis’ own failure to heed his warnings and Davis’ failure to support Memminger before Congress.56

After several years, it became clear that the heavily debt- and note-dependent strategy for financing the war was untenable. While there were a few holdouts, even some of the most ardent of states’ rights supporters pleaded to be taxed. Governor Joseph Brown of Georgia, a states-righter who had quarreled with President Jefferson Davis over something so basic as the control of Georgia troops in the prosecution of the war, exclaimed "[f]or God's sake tax us. Nothing else can save us from ruin." As Assistant Secretary of State William M. Browne lamented, “[i]f something is not done, we shall have to pay $27,000 for a barrel of flour.” Moreover, the alternative, which consisted of the continued impressment of goods and property by the Confederate army,

55 Quoted in Seligman, supra note 34, at 483.
56 Ball, supra note 28, at 8. On several occasions, Memminger personally warned Davis of the inflationary risks posed by the continued issuance of Treasury notes. Lerner, supra note 37, at 520.
57 The Wilmington Journal wrote in the spring of 1863: “Let our authorities then fearlessly task and stretch the public credit to the utmost, in order to carry on the war so that taxation may not crush to earth our already overburdened people.” Lerner, supra note 37, at 509, n.22.
58 See Razaghian, supra note 27, at 30 (finding, from a roll call analysis of the Confederate Congress, that "as hostilities stretched out into a protracted war, strategies adjusted accordingly. By the Winter of 1863, Congressmen from the Deep South were more likely to support taxes in the Confederacy's financial portfolio than other legislators.").
60 Quoted in Paul, supra note 3, at 19.
61 Quoted in Ball, supra note 28, at 231.
was arguably worse because of its inherent arbitrariness. The Confederacy thus began to consider a more comprehensive financial policy—one that included an income tax and other internal taxes.

On April 24, 1863, a year behind the Union and more than four months after Lincoln issued the Emancipation Proclamation, the Confederate Congress passed a bill levying a graduated income tax with a higher rate on unearned than earned income as well as a tax on corporate income. This attempt to reach “idle wealth” may in part have been a reaction to the ability of the wealthy to evade contributing to the effort through conscription in the military. Only a year earlier, the Confederate Congress enacted the first draft in American history to counter the waves of soldiers who failed to re-enlist after their year of service was up. All able-bodied white males between ages eighteen and thirty-five were required to serve, but men could avoid service either by falling into one of the necessary classes of exempted occupations, such as overseers on large plantations, or by hiring a substitute from outside the pool of those already required to serve. By late 1863, rich men paid as much as $300 worth in gold to hire a substitute

---

62 Id. at 231. Impressment occurred when the Army took items and paid for them at a below market rate that effectively constituted a tax on the owners. Farmers eventually stopped producing as much since they knew their crops would be taken by the Army at a fraction of their real value once they tried to bring them to market.

63 For unearned income, the lowest rate was five percent on income between $500 and $1,500 and the highest rate was fifteen percent on incomes in excess of $10,000. For earned income, there was a one percent tax on income between $1,000 and $2,000 and a two percent rate on the amount in excess of $2,000. James M. Mathews, ed. Public Laws of the Confederate States of America, Passed at the Third Session of the First Congress, 1863 (Richmond: R.M. Smith, 1863): Sec. 8(VI) (hereinafter Confederate Statutes). Ball, supra note 28, at 233. Corporations were required to reserve one-tenth of their annual earnings, set apart for dividends and reserves, and remit that to the Collector. Any dividend subsequently paid would thus be exempt from further tax. Where the earnings exceeded ten percent of the company's capital stock, the rate of tax increased from one-tenth to one-eighth and where the earnings exceeded twenty percent of the company's capital stock, the rate of tax increased from one-eighth to one-sixth. Confederate Statutes, supra, Sec. 8(VI); Steven A. Bank, "Entity Theory as Myth in the Origins of the Corporate Income Tax," William & Mary Law Review 43(2) (December 2001): 447, 488-89

64 McPherson, supra note 13, at 430.
soldier, an amount equal to as much as three years’ worth of wages for a skilled man.\textsuperscript{65} The combined effect of the exemptions and the ability to substitute led some to call the conflict “a rich man’s war but a poor man’s fight.”\textsuperscript{66} To partially alleviate this concern, the Confederate Congress went so far as to impose a special tax on exempted overseers in which the revenues were disbursed to the families of Confederate soldiers.\textsuperscript{67}

In addition to income taxes, the 1863 revenue bill imposed a one year ad valorem tax on agricultural and luxury products, a tax on currency and bank notes, license taxes on a variety of professions, ranging from bankers and lawyers to innkeepers and entertainers, a tax on gross sales, and a profits tax on speculative commodity transactions.\textsuperscript{68} It also imposed an in-kind tax on agricultural products grown in 1863, sometimes referred to as a “tithe tax” in order to institutionalize and reform the impressments system.\textsuperscript{69} Notably absent from this panoply of taxes were any property taxes on real estate, slaves, or livestock, which were excluded because of constitutional concerns.\textsuperscript{70} Unlike the Provisional Constitution in effect when the 1861 direct taxes were enacted, the new constitution required apportionment and the majority concluded that the lack of an adequate census made this task a practical impossibility.\textsuperscript{71}

While the 1863 tax was a large step forward compared with the Confederacy’s earlier financial path, it faced substantial obstacles. Secretary Memminger had warned that licenses and other similar taxes were objectionable because they required “vexatious

\textsuperscript{65} Id. at 432.
\textsuperscript{66} Id. at 431.
\textsuperscript{67} Bensel, supra note 47, at 183. The Confederacy also eventually repealed substitution altogether by the end of 1863. McPherson, supra note 13, at 603.
\textsuperscript{68} \textit{Confederate Statutes}, supra note 62, at Secs. 1, 5; Ball, supra note 28, at 233-34; Bensel, supra note 47, at 169.
\textsuperscript{69} \textit{Confederate Statutes}, supra note 62, at Sec. 11; Bensel, supra note 47, at 169; Ball, supra note 28, at 235.
\textsuperscript{70} Ball, supra note 28, at 234.
\textsuperscript{71} Id. at 215-216, 234.
and expensive machinery."\textsuperscript{72} It soon became apparent that "the novelty of many of the provisions of the law had given rise to serious objections."\textsuperscript{73} More seriously, the Confederacy simply did not have the administrative capacity to accommodate a tax bill covering so many subjects and in such complexity. As Secretary Memminger later reported,

the difficulties which are encountered in the collection can only be estimated by any one who will inspect the mass of papers which are required for each return, and the inquiries necessary to be made of each individual taxpayer. The results of the tax will probably confirm the recommendation already made of a resort to a more simple system of taxation. The frauds and evasions, which cannot be discovered under the present system, are a perpetual drain upon the tax, which is necessarily increased by the number of officers who must be employed in it collection\textsuperscript{74}

Not surprisingly, therefore, collections were initially quite slow. Almost a year after passage, only $27 million had been received, although that number excluded the value of in-kind payments.\textsuperscript{75} Collections eventually rose to more respectable levels as the Confederacy better developed its administrative apparatus,\textsuperscript{76} but they were still well below expected levels.\textsuperscript{77} To augment revenues, the Confederate Congress increased the rates on income when the tax act was extended on February 17, 1864.\textsuperscript{78} This came too late in the war, however, to have much effect. Memminger reported in May 2, 1864 that the Confederacy could only be expected to collect in-kind tax revenues during the

\textsuperscript{72} Lerner, \textit{supra} note 37, at 510.
\textsuperscript{73} Seligman, \textit{supra} note 34, at 490.
\textsuperscript{74} Quoted in Seligman, \textit{supra} note 34, at 492.
\textsuperscript{75} Ball, \textit{supra} note 28, at 236; Bensel, \textit{supra} note 47, at 170.
\textsuperscript{76} Bensel, \textit{supra} note 47, at 170. By 1864, the war taxes raised almost eleven percent of the Confederacy’s total revenues, a figure that compares favorably with the fifteen percent raised by taxation in the North during the same period. Schwab, \textit{supra} note 31, at 41.
\textsuperscript{77} See Razaghian, \textit{supra} note 27, at 38 (Table 8). In all but two states -- North Carolina and Virginia -- the taxes collected fell far short of what was expected based on the value of taxable property.
coming year.\(^{79}\) Surrender occurred before the 1864 taxes could ever be fully implemented.\(^{80}\)

Although the need for fiscal sacrifice was readily apparent and in large part unfulfilled,\(^{81}\) Confederate taxes were still perceived as exceedingly burdensome. After the April 1863 tax act was passed, one southern partnership posted an advertisement in the Richmond \textit{Whig} explaining that "[i]n consequence of the very heavy taxes, we have deemed it best to close our business."\(^{82}\) The \textit{Richmond Enquirer} reported that "[t]he tax bill . . . carries consternation to the whole mercantile community. Half the merchants and manufacturers talk of discontinuing their business."\(^{83}\) Moreover, the protest was not just from the urban business community. Farmers objected on the ground that the income tax was payable in depreciated Confederate currency while the in-kind tax necessarily was paid at full market value.\(^{84}\) They organized protests and passed resolutions denouncing the tax act as “unjust and tyrannical,” “monarchical,” and “a relic of barbarism which alone is practiced in the worst despotisms.”\(^{85}\) Secretary Memminger even conceded that the system was "so cumbersome and intricate that delay and disappointment will be its inevitable result."\(^{86}\) It may be that many of the Confederacy’s financial difficulties were simply unsolvable,\(^{87}\) but it is clear that at least in the tax context this had something to do

\(^{79}\) Seligman, \textit{supra} note 34, at 491.
\(^{80}\) See Ratner, \textit{supra} note 40, at 107, Paul, \textit{supra} note 3, at 22.
\(^{81}\) Over the course of the war, only about one percent of the Confederacy’s cash receipts were raised through taxation, which is a shockingly small proportion in comparison to modern wartime governments. David M. Potter, "Jefferson Davis and the Political Factors in Confederate Defeat," in \textit{Why the North Won the Civil War} (Baton Rouge: Louisiana State University Press, 1960): 94.
\(^{82}\) "The New Rebel Tax Law," \textit{N.Y. Times}, May 16, 1863 (reprinting an advertisement from the Richmond \textit{Whig}).
\(^{83}\) "Affairs in Richmond," \textit{N.Y. Times}, April 26, 1863, at 2 (reprinted from the Richmond Enquirer, April 16, 1863).
\(^{84}\) Seligman, \textit{supra} note 34, at 489; Paul \textit{supra} note 3, at 20.
\(^{85}\) Id.
\(^{86}\) Quoted in Paul, \textit{supra} note 3, at 21.
\(^{87}\) Bensel, \textit{supra} note 47, at 172.
with the initial, and to a large extent continuing, unwillingness to sacrifice the principle of decentralization where it concerned taxation.

B. Union

At the beginning of the war, the Union appeared to suffer from the same excess of optimism that plagued the Confederacy. Salmon Chase, Lincoln's Secretary of the Treasury, proposed to reduce the massive deficit through a combination of traditional means, including issuing Treasury notes with the assistance of financier Jay Cooke,\(^88\) increasing tariffs to their pre-1857 levels, imposing or raising excise taxes, instituting license fees, and selling public lands.\(^89\) This was consistent with Chase's own belief, and the belief of many others, that the war would be of short duration and therefore extraordinary fiscal measures were either unnecessary or would be temporary: "[i]t is hardly to be doubted, moreover, that the great body of the citizens . . . will, ere long, become satisfied that order and peace . . . are preferable to disorder and conflict. . . It is not unreasonable to expect that with restored Union will come . . . renewed prosperity."

The New York Times offered a similar picture of the mood in Northern commercial circles, reporting that "[t]he belief is almost universal, among our business men, that in a short time the whole thirty-four States will be, in practical truth, as they are by the letter and spirit of the Constitution, One People."\(^91\)

---

\(^88\) The loan strategy began even before the hostilities began with the firing on Fort Sumter on April 12. On March 22, Chase opened bids on government securities totaling $8 million at a six percent rate. Smith, supra note 7, at 10. On July 17 and August 5, 1861, Congress loan acts that empowered the Secretary to borrow $250 million, using a combination of three-year Treasury notes at a rate of 7.3 percent or twenty-year bonds with a rate not to exceed 7 percent. Ratner, supra note 40, at 64.

\(^89\) Witte, supra note 25, at 67.

\(^90\) Quoted in Smith, supra note 7, at 15-16.

\(^91\) "Monetary Affairs," N.Y. Times, July 18, 1861, at 3. The New York Tribune, reporting the following day on the advances of the Union army in Virginia, wrote "[i]t seems to be the universal and joyful conviction that this is the 'beginning of the end.'" Quoted in Smith, supra note 7, at 16.
This optimism took a significant hit with the Confederate victory in the first Battle of Bull Run (or Manassas, as the South called it) on July 21, 1861. Because of the battle’s proximity to Washington, D.C., Congress was intimately aware of the defeat. It could hardly be otherwise considering that wild-eyed Union soldiers retreated so far from the battle that they ended up back in the capitol where members of the House and Senate had to go out to the streets themselves to try to persuade them to turn back and fight. Congress also was well aware of the defeat’s implications for Union finances. Republican Senator James F. Simmons, a manufacturer from Rhode Island, commented "I do not know what to call the sort of precipitate retreat from Centreville that some of our soldiers made the other night; but I believe it is going to make it much more difficult for us to negotiate loans than if they had stood their ground." At this point, morale in the North was at its lowest point, but it was perhaps the prod it needed. As the London Times correspondent predicted, “[t]his prick in the great Northern balloon will let out a quantity of poisonous gas, and rouse the people to a sense of the nature of the conflict on which they have entered.”

This turning point in the war may have been a turning point in fiscal matters as well. On July 23, 1861, two days after the disastrous defeat at Bull Run, Thaddeus Stevens, the "dictatorial" Republican chair of the House Ways and Means Committee, proposed a $30 million direct tax on land, with each state's share to be apportioned by population. Two days later, Senator Simmons proposed adding an income tax provision.

---

92 The battle took place only a day’s march from Washington, D.C. Bensel, supra note 47, at 177 n. 171.
93 McPherson, supra note 13, at 345.
94 Con. Globe, 37th Cong., 1st Sess. 254 (July 25, 1861) (statement of Sen. Simmons, R-R.I.). Centreville was a town was just to the east of the fighting and twenty miles west of Washington, D.C. See McPherson, supra note 13, at 343.
95 Quoted in McPherson, supra note 13, at 348.
96 Ratner, supra note 40, at 64.
to the tariff bill passed in the House on July 18.\textsuperscript{97} The consensus appeared to be that relying on loans to finance the war would be insufficient. When Stevens introduced his proposal, the \textit{New York Times} wrote "[i]t is admitted on all hands that the money to be raised on loan, by the Government during the present fiscal year, will fall short of its actual necessities, by at least eighty millions -- and that taxation, either direct or indirect, or both, must be resorted to, without delay, to provide for the deficiency."\textsuperscript{98} The consensus seemed to be that the war could no longer be prosecuted without demanding monetary sacrifices from its people in the form of higher taxation. Stevens argued that "the annihilation of this government is the alternative."\textsuperscript{99} Others agreed:

Hitherto we have been able to maintain a powerful government at comparatively a small expense raised by indirect taxation; none felt the burden of the government, while all enjoyed its inestimable blessings. . . . [N]o one could have conceived the ridiculous idea of a government undertaking to sustain itself from overthrow against a formidable rebellion, without a most serious drain upon the taxable resources of the people. The cost of our vast military and naval operations necessary to achieve the great ends which the government has in view are enormous, and it rests with the people now to say whether these ends shall be urged forward or whether we shall let the government go to a quiet and ignominious grave.\textsuperscript{100}

Stevens' proposal proved controversial, however, since farmers would bear the brunt of the tax while holders of intangible wealth such as stocks and bonds would be virtually exempt.\textsuperscript{101} Thus, the House initially recommitted the bill with instructions to amend it to include a tax on both real and personal estate. When this ran into constitutional

\textsuperscript{98} "The War Tax," \textit{N.Y. Times}, July 23, 1861, at 4. The correspondent from the London Times concurred, noting that there are some in Europe "who will hesitate to trust in time of war, and for purposes of unproductive investment, a Federation of States, several of which in their individual capacity repudiated a debt in time of peace and prosperity." "Comments on the Budget of Secretary Chase," \textit{London Times}, in \textit{N.Y. Times}, August 2, 1861, at 3.
\textsuperscript{99} Quoted in Ratner, \textit{supra} note 40, at 64.
\textsuperscript{100} "Remarks on the Direct Tax Bill," \textit{Scientific American} V(7) (August 17, 1861): 105.
difficulties, the bill was again recommitted with instructions to replace much of the direct tax portion with an income tax. At the end of July, both the House and Senate passed bills that, to varying degrees, combined an income tax with a direct tax on land. A compromise bill was adopted by the full Congress on August 5, 1861.102

As in the Confederacy, an immediate obstacle to the new direct tax was the absence of a federal system for collecting the new direct tax. There had been concern that the Act "would send federal tax assessors and collectors into all parts of the country." 103 To help alleviate this concern, Congress sought to piggyback on state tax collection systems. Unlike in the Confederacy, however, states were not required to collect a certain quota of taxes due in their state, but rather were only asked to voluntarily assume this task.104 Indeed, they were provided an incentive in the form of a reduction in their state’s share if they agreed to collect the taxes.105 Notwithstanding this mechanism for enlisting the support of state collection systems, the 1861 Act also provided for the erection of a Federal system of internal revenue collections. The Act specified the duties of assessors and described the penalties for non-compliance.106 Additionally, the President was authorized, upon the nomination of the Secretary, to appoint an officer in the Treasury to serve as a Commissioner of Taxes, to divide the country into collection districts, and to appoint an assessor and collector for each district.107

102 Act of August 5, 1861, 12 Stat. 297. The House bill had proposed a three percent rate on all incomes over $600 and the Senate had proposed a five percent rate on all incomes over $1000. In Conference, the houses settled on a three percent rate on income in excess of $800. Witte, supra note 25, at 68.
103 Smith, supra note 7, at 24-25.
105 Smith, supra note 7, at 25. The state could secure a fifteen percent reduction if they assumed payment by February and remitted by the end of June and a ten percent reduction if they remitted by the end of September.
While the 1861 Act appeared to promise a radical change, it initially did little to improve finances. The direct tax produced no revenues since assessments did not have to be made until April of 1862. Moreover, all of the states but Delaware elected to assume the tax obligation themselves, making it little more effective than the Confederate version of this tax. The income tax fared even worse, as Chase made no effort to assess or collect it. He also failed to nominate anyone to serve as Commissioner of Taxes, further frustrating the erection of any collection effort on internal taxes.

Unlike in the Confederacy, where the relatively unsuccessful 1861 direct tax remained on the books for two years, the Union appeared to more quickly appreciate the gravity of its financial situation. War expenses were steeply rising and the national debt was rising at a rate of two million dollars a day. By January of 1862, the ability to secure loans of any kind was severely hampered by the decision of the government and banks throughout the country to suspend specie payments. Chase modestly proposed raising $50 million in additional revenues, but the House Way and Means Committee recognized this would not be enough and called for the imposition of a tax that would raise three times that amount. Rather than immediately proposing the taxes it would employ to procure this sum, the Committee proceeded deliberately to craft a bill, often to

109 Bensel, supra note 47, at 168. A primary difference was that while in the Confederacy most states paid their quota through issuances of state treasury notes, which compounded the inflation effect, in the Union at least nine of the states, plus the District of Columbia, chose to pay their quota by increasing their own general property tax. Smith, supra note 7, at 38-39. Only five states met their quotas through borrowing. Id. at 39.
110 Ratner, supra note 40, at 70; Smith, supra note 7, at 25.
111 Smith, supra note 8, at 24, n. 4.
112 Howe, supra note 14, at 54. By the spring of 1862, the debt was at $505 million. Witte, supra note 25, at 69.
113 Ratner, supra note 40, at 70.
114 "How the Expenses of the War are to be Met," N.Y. Times, January 8, 1862, at 1.
the frustration of the popular press. When it first introduced a bill in March of 1862, the Saturday Evening Post called it "one of the longest of any kind ever before prepared - months of preparation having been bestowed upon it." 

This new bill dramatically expanded the base of taxation, most notably in the arena of internal excise and gross receipts taxes. A contemporary observer from Austria commented that under the bill, "[t]he citizen of the Union paid a tax every hour of the day, either directly or indirectly, for each act of his life; for his movable and immovable property; for his income as well as his expenditure; for his business as well as his pleasure." In addition to the direct tax and a somewhat broadened version of the income tax adopted in 1861, the bill included (1) excise taxes on everything from alcohol and tobacco to soap and salt; (2) license taxes on everything from bowling alleys and circuses to bankers and commercial brokers; (3) stamp taxes on the full range of legal and commercial documents; and (4) both an inheritance tax and a gross receipts tax on corporations in the transportation and financial industries. Perhaps the most controversial aspect of the new tax bill was a general three percent excise tax on all manufactured goods, which was expected to raise more revenue from any single item

---

117 Howe, supra note 14, at 65. Several popular journals made light of the breadth of the new internal excise taxes. The Knickerbocker Magazine proposed several humorous additions to the bill, including a $0.25 cent tax for speaking French or sneezing "with unusual noise." “Additions Proposed to the Tax Bill,” Saturday Evening Post, July 26, 1862, at 6 (reprinted from Knickerbocker Magazine). Scientific American observed that the committee had thankfully omitted crinoline, a component for the production of “hooped skirts,” from the tax list: "The fair portion of the community need not dread any retrenchment of their circumference; the gallantry of the Congressional committee has spared them this mortification.” “The Tax Bill Amendment,” Scientific American VIII(12) (March 21, 1863): 177.
118 It proposed a three percent tax on incomes in excess of $600 (which was the same as the House bill passed in 1861), as compared with the 1861 Act's three percent rate on incomes in excess of $800. "The Tax Bill," Saturday Evening Post, March 15, 1862, at 3.
119 Id.; Stanley, supra note 1, at 29.
except customs duties. Manufacturers argued that this tax, which was based on sales rather than profits, would be “utterly ruinous to them” and would lead to “triple and quadruple taxation” because it was imposed in addition to the excise taxes on raw materials. To placate manufacturers and to avoid "destroy[ing] the goose that lays the golden eggs," Representative Justin Morrill proposed to introduce a bill for an increase in tariff rates as a companion to the tax bill. With the memory of the Union's victory at Shiloh a few days earlier fresh in their mind, and perhaps in the hopes of driving a stake through the rebellion only a few months after Jefferson Davis' inauguration as President of the Confederate States of America, the House passed the bill.

In the Senate, the direct tax portion of the bill was replaced with a graduated version of the income tax. The theory was that, because of the constitutional requirement of apportionment, the direct tax on land would unduly burden landholding farmers in sparsely populated states as compared with urban dwellers who primarily held intangible wealth such as stocks and bonds. While the progressive feature of the income tax served to replace the revenue from the direct tax, it also operated to balance out the more regressive effects of the excise taxes and tariff duties, which principally burdened consumption. As Senator Simmons explained, "[t]he poor people pay as much, and

---

120 “The Tax Bill,” *Saturday Evening Post*, March 22, 1862, at 7. Customs duties were expected to raise $50 million and the excise tax on manufactured goods was expected to raise $30 million. The next largest item was the tax on spirits, which was expected to raise $15 million. Id.
122 Howe, *supra* note 14, at 58; Stanley, *supra* note 1, at 29 (quoting Representative Justin Morrill).
124 Under the proposal, the income tax would have had a three percent rate on incomes between $600 and $10,000, a five percent rate on incomes between $10,000 and $50,000 and a seven-and-a-half percent rate on incomes in excess of $50,000. Cong. Globe, 37th Cong., 2d Sess. 2486 (1862) (statement of Sen. Simmons).
125 See Bank, *supra* note 5, at 346-47.
rather more, generally, than rich men on their consumption." The income tax also contained some innovative features, such as a reduced rate on interest from Federal securities to encourage their purchase, a higher rate on income from property in the U.S. owned by citizens living outside the country on the theory that they would not otherwise contribute through the internal excise taxes, and withholding tax provisions for dividends, interest, and government salaries. After an adjustment of the rates of progression in Conference Committee, plus an agreement to suspend the direct tax for two years rather than eliminate it altogether, Congress passed the bill and President Lincoln signed it into law on July 1, 1862.

The most significant distinction between the 1862 Act and both the Union and Confederate acts from the previous year may have been the institutional development that it engendered. Implementation was a significant concern when the Act was passed. The New York Times predicted that the “very multiplicity of the objects and persons subject to the new impost . . . will render the administration of the law both delicate and arduous.” To handle this new task, the Office of Internal Revenue was created under the auspices of the Treasury Department. With three clerks borrowed from other departments of the Treasury, its first Commissioner, George S. Boutwell, drafted all of the regulations and issued the rulings that interpreted the new Act. Its direct influence,

---

127 See Hill, supra note 2, at 426.
128 The final bill eliminated the top rate, applying a five percent rate on all income above $10,000.
129 Act of July 1, 1862, ch. 119, 12 Stat. 432, 473.
131 Ratner, supra note 40, at 75. A Office of Internal Revenue was created for the first time in 1792 before being dismantled in 1802. It was revived in 1813 to administer the internal taxes enacted in the War of 1812. It has existed continuously since 1862, now operating under its modern title of the Internal Revenue Service. Giroux, supra note 4, at n. 1.
132 Id. at 76.
however, was felt via the hundreds of assessors and collectors dispersed throughout the Union's 185 collection districts to implement the 1862 Act's provisions.\textsuperscript{133}

One immediate concern was training and preparation. The Act was supposed to go into effect on August 1, but was delayed for a month in order to permit the Office to put its organizational structure into effect.\textsuperscript{134} This was apparently not enough time. When the Act finally became effective on September 1, most appointments of assessors and collectors had yet to be confirmed or trained. "In fact," it was reported, "the greater number of these gentlemen are quite in the dark as to their future proceedings, and the questions of anxious inquirers elicit among them only a helpless stare of perplexity."\textsuperscript{135}

Given the comprehensive nature of the Act and the complexity of the task of implementation, it was not surprising that the new Office of Internal Revenue encountered resistance. There were familiar concerns expressed over the centralization of government power through the creation of the Internal Revenue Department and the designation of collection districts,\textsuperscript{136} but the most serious threat came from the manufacturing community. One letter writer reported in December that "[n]ot one manufacturer in a thousand has paid his ad valorem taxes; and this part of the law threatens to become a nullity and yield comparatively no income."\textsuperscript{137} At a convention in Chicago held in June of 1863, manufacturers went so far as to develop a circular "urging upon the Commissioner of Internal Revenue to either ignore the law altogether, or to

\textsuperscript{133} Bensel, supra note 47, at 169.
\textsuperscript{137} Letter to the Editor, "Revision of the Tax Bill," \textit{N.Y. Times}, December 14, 1862, at 3.
suspend its operation until Congress can meet and repeal the obnoxious feature of the bill."  

This was a crucial test to the legitimacy of the Office of Internal Revenue and it reflects the fact that patriotic fervor and the calls for sacrifice could not overwhelm the self-interest that characterizes modern tax politics. The Merchants Magazine and Commercial Review noted that "[i]f [manufacturers] seek to evade it, it is to be feared their example may have a bad influence upon other classes, less interested, perhaps, in sustaining the government." What could have been a serious problem, however, never came to pass. The Office responded to the Chicago convention resolutions firmly: "Rest assured that the law is not considered by this office in the light either of a mistake or an accident, and that its provisions will be neither explained away, nor its operations suspended." The Office was helped in its efforts by the inequity of the manufacturers' claims, especially in light of the sacrifices borne by others in the same act. One journal wrote "[t]he tax on manufacturers is doubtless heavy, and will fall with great inconvenience on many; but we, as publishers, might as well claim exemption from the onerous burthen as the manufacturers. We pay a heavy tax on paper, the price of which has greatly increased since the war, also upon every other article used in our business."

---

141 “The Tax on Manufacturers,” Scientific American VIII(25) (June 20, 1863): 393. One attorney had a similar response, writing that the tax, "both specific and ad valorem, is invariably added to the prime cost of the manufactured article, and ultimately falls either wholly upon the consumer, or is partially shared by the retailer . . . The income tax complained of is really the only government charge in any shape which
The manufacturers appeared to be arguing for special treatment, which did not help their case. Some manufacturers may have recognized the unpopularity of their pleas, especially in light of their existing tariff protection.

Notwithstanding the monumental achievement in establishing the Office of Internal Revenue in such a short period of time, administrative difficulties clearly hampered the collection process. For instance, most of the stamp taxes still could not be enforced several months after the Act's effective date because of the delay in preparing the stamps themselves.\textsuperscript{142} The definition of income was also unclear, with the Act silent on such crucial questions as the taxation of capital gains from the sale of real estate.\textsuperscript{143}

Coupled with the decline in tariff revenues from the activities of the Confederate navy and the pressures of inflation caused by the heavy issuances of Treasury notes,\textsuperscript{144} there was once again a call for further taxation. The \textit{New York Times} demanded "to be taxed to a degree which shall fairly correspond with the vast amount of promissory money afloat."\textsuperscript{145} Despite the existing heavy burden, the popular wisdom was that the public could bear it. The \textit{Times} opined that "[o]ne of the greatest fallacies into which many are beguiled is, that the imposition of taxes in addition to those already imposed will bear too

\textsuperscript{142} "The Stamp Act," \textit{N.Y. Times}, October 31, 1862, at 2.
\textsuperscript{143} See Hill, \textit{supra} note 2, at 430. The Commissioner ruled that all gains from the sale of real estate must be included in income during the year in which the property was sold. In the 1864 Act, Congress specifically provided that only profits arising from the sale of property purchased in the same taxable year must be included. The 1867 Act extended that rule to profits arising from sales of property purchased within the last two years.
\textsuperscript{144} Ratner, \textit{supra} note 40, at 82.
\textsuperscript{145} Quoted in id. at 90.
heavily on the people.\textsuperscript{146} The \textit{New York Tribune} echoed this point, noting "the reports thus far . . . show an unexpected capacity of our people for taxation.\textsuperscript{147}

Thus, in two revenue acts adopted in 1864, Congress increased the tax rates to what were described as "extravagant" levels.\textsuperscript{148} Internal excise tax rates were more than doubled on spirits, tobacco, and cigars and the general excise taxes increased correspondingly.\textsuperscript{149} Income tax rates increased as well, with the highest rate topping out at ten percent on income over $10,000 under the Act passed on June 30, 1864.\textsuperscript{150} A second emergency income tax act passed on July 4, 1864 increased all rates by another five percent on incomes above $600 to cover bounties necessary to recruit and retain soldiers.\textsuperscript{151}

Much of the debate over the income tax surrounded the equity of introducing steeper progressivity to the system.\textsuperscript{152} Part of the rationale for progressive rates was pragmatic—the need for revenue and the appearance of fiscal stability. Augustus Frank, a Republican railroad director from upstate New York argued for a more progressive income tax on the grounds that it was necessary to reassure private creditors: "the larger the tax we pay at this time, the safer we are and the better will be the securities of the Government."\textsuperscript{153} Representative J.B. Grinnell, a Republican from Iowa concurred in this sentiment, explaining “if gentlemen demand to sustain our credit and prosecute the war

\textsuperscript{146} "Popular Fallacies Regarding Taxation," \textit{N.Y. Times}, April 8, 1864, at 4.
\textsuperscript{147} Quoted in "Chapter on Taxation: Our Capacity for Taxation," \textit{Advocate of Peace} 15 (January/February 1864): 24.
\textsuperscript{148} Howe, \textit{supra} note 14, at 62.
\textsuperscript{149} Ratner, \textit{supra} note 40, at 82 n.9.
\textsuperscript{150} Act of June 30, 1864, ch. 173, 13 Stat. 223.
\textsuperscript{151} Act of July 4, 1864, 13 Stat. 417; Ratner, \textit{supra} note 40, at 89.
\textsuperscript{152} Cong. Globe, 38th Cong., 1st Sess. 1876-77, 2513-15 (April 26 & May 27, 1864); Hill, \textit{supra} note 2, at 433.
\textsuperscript{153} Id. at 1876 (April 26, 1864).
with success, I could not advocate anything else in justice to the middle classes in this country.\textsuperscript{154}

More generally, the income tax and its progressive rates were justified as a relatively small but importance sacrifice for the wealthy in light of the sacrifices made on their behalf on the battlefield. Representative Grinnell made a statement on the floor that typified this view:

\begin{quote}
It is time that extravagance in gewgaws, snobbishness in display, and that large class whose great care is to safely compound their hundreds of thousands, should feel that there is a war and a demand which they have not yet felt on their purses and on their patriotism, which in many cases has been but poorly acknowledged. . . . Let colossal wealth, protected, meet the full share of burdens, which are lighter in this country than in any other civilized country in time of war.\textsuperscript{155}
\end{quote}

The issue was not simply that the rich had not yet been forced to make material sacrifices, but that they were not sacrificing their lives and the lives of their sons in battle. The Union had followed the Confederacy in enacting a draft, but men of means could not only avoid service by locating a substitute, as was true in the Confederacy, but for most of the war they could also pay a $300 fee to the government to commute the service obligation until the next draft.\textsuperscript{156} The combination of substitution and the implicit commutation tax meant that few wealthy men served and even when they did it was as an officer rather than on the front lines.\textsuperscript{157} One newspaper declared that "[t]he rich are

\begin{footnotes}
\item[154] Id.
\item[155] Id. at 1876-77.
\item[156] McPherson, supra note 13, at 601. Commutation was in part designed to avoid the unsavory tactics of "substitute brokers" by allowing the government to find a substitute for a fee far below that charged by most brokers. "General Intelligence," Christian Advocate and Journal, July 23, 1863, at 30; "The Alleviations of the Draft," The Independent -- Devoted to the Consideration of Politics, Social, and Econ. . . July 23, 1863 at 764. Nevertheless, commutation was demonized by its opponents and was eventually repealed for all but conscientious objectors in the middle of 1864. McPherson, supra note 13, at 601
\item[157] Between substitution and commutation, most men of even moderate means and resources could avoid conscription. While 207,000 men were drafted, 87,000 paid a fee to have their service obligation commuted and 74,000 paid substitutes to take their place. Id.
\end{footnotes}
exempt!"158 This was used to justify even greater progressivity in the income tax. Senator Garrett Davis of Kentucky employed such logic in arguing against a proposal that would remove the ten percent rate on incomes in excess of $25,000:

The idea that millionaires and men whose incomes exceed $25,000 as a general rule go into the camp is not supposable. There may be some exceptional cases; but if they go into camp at all it is not by shouldering the musket, unless in very rare cases. They do not send their sons there as a general rule unless the sons go with epaulets upon their shoulders.159

The implication was that the higher tax burden on wealthy individuals was justifiable given they were not otherwise sacrificing for the cause.

Although the 1864 Acts produced some of the highest receipts during the war, bond sales began to lag and President Lincoln abruptly accepted Secretary Chase's resignation. His successor as Treasury Secretary, former Senate Finance Committee chair William Pitt Fessenden, engaged Jay Cooke to market several more loan issuances.160 It was not enough, however, to finance the final military push. Thus, in December of 1864, Fessenden recommended once again increasing internal revenue taxes with the goal of raising nearly $300 million.161 He hoped to satisfy much of this need with increased progressive income taxation, proposing to eliminate exemptions and thereby broaden the base.162 The increasing reliance on the income tax was particularly controversial, especially in light of the detested custom of publicizing returns to aid

---

158 Id. at 602. This charge may have been unfair. According to several studies, there was little correlation between who paid the commutation fee and wealth, in part because some local governments paid the fee on behalf of their residents out of property tax revenues and some employers did the same to retain workers. Id. at 602-604.
160 Ratner, supra note 40, at 96.
161 Id.
162 Id.
collection, but it was by no means the only objectionable aspect of the new revenue bill introduced by Justin Morrill on February 9, 1865. As Frederic Howe observed, "every possible article which had escaped the scrutiny of the earlier Acts was sought out and taxed." Nevertheless, the measure was eventually adopted on March 3, 1865.

As a result of the 1864 and 1865 Acts, the percentage of tax revenues derived from internal taxes increased substantially, with income tax receipts more than tripling.

More broadly, as Table 1 illustrates, these Acts marked the culmination of what can only be described as a radical shift in the system of Federal financing over a very short period.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Customs</th>
<th>Income</th>
<th>Other Internal Revenue</th>
<th>Other*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1866</td>
<td>558,032,620</td>
<td>179,046,652</td>
<td>72,982,159</td>
<td>236,244,654</td>
<td>69,759,155</td>
</tr>
<tr>
<td>1865</td>
<td>333,714,605</td>
<td>84,928,261</td>
<td>60,979,329</td>
<td>148,484,886</td>
<td>39,322,129</td>
</tr>
<tr>
<td>1864</td>
<td>264,626,771</td>
<td>102,316,153</td>
<td>20,294,732</td>
<td>89,446,402</td>
<td>52,569,484</td>
</tr>
<tr>
<td>1863</td>
<td>112,697,291</td>
<td>69,059,642</td>
<td>2,741,858</td>
<td>34,898,930</td>
<td>5,996,861</td>
</tr>
<tr>
<td>1862</td>
<td>51,987,456</td>
<td>49,056,398</td>
<td>N/A</td>
<td>N/A</td>
<td>2,931,058</td>
</tr>
<tr>
<td>1861</td>
<td>41,509,931</td>
<td>39,582,126*</td>
<td>N/A</td>
<td>N/A</td>
<td>1,927,805</td>
</tr>
</tbody>
</table>

**Table 1.**

**Federal Receipts: 1861-1866**

---

163 Assessors and collectors were authorized to publicize self-reported income levels in an attempt to ferret out fraud and underreporting. This was a widely disliked practice. See, e.g., Carl Benson, Letter to the Editor, “Income Tax Returns,” *N.Y. Times*, Feb. 5, 1865, at 2 (“The authorized publication of the Income Tax lists at first struck every man who had any feelings of his own or any respect for those of others as a most disagreeable and injudicious interference with private affairs.”); Hill, *supra* note 2, at 436.

164 Ratner, *supra* note 40, at 97. Under Morrill's proposal, the top rate would increase from 7.5 percent to 10 percent on incomes over $3,000. The House eventually settled on a more modest increase to a tax of ten percent on incomes over $5,000.

165 Howe, *supra* note 14, at 68.

166 Act of March 3, 1865, ch. 78, 13 Stat. 469. The 1865 Act's income tax feature was later subject to constitutional challenge as an unapportioned direct tax. The Court upheld the tax on the grounds that it was more similar to a duty or excise than to a direct tax. Springer v. United States, 102 U.S. 586 (1880).

167 The raw numbers are derived from Historical Statistics, *supra* note 16, at 297, Series P 89-98. The percentages are calculated based on these raw numbers. Because the percentages have been rounded up, they do not always add up to one hundred.

168 This category primarily includes the proceeds from sales of public lands and the sales of government-owned securities. Id.
While customs revenues as a percentage of total receipts declined precipitously during the war, other revenue sources very quickly filled the gap. Internal excise taxes made up the large bulk of the new revenues, but income taxes also became a significant part of the revenue system.

Not surprisingly, until fighting ceased and the soldiers went home, this increase in the size of revenues never quite kept pace with the even more significant rise in government expenditures depicted in Table 2:

Table 2.170

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Expenditures</th>
<th>Total Receipts</th>
<th>Surplus/(Deficit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1866</td>
<td>520,809,417</td>
<td>558,032,620</td>
<td>37,223,203</td>
</tr>
<tr>
<td>1865</td>
<td>1,297,555,224</td>
<td>333,714,605</td>
<td>(963,840,619)</td>
</tr>
<tr>
<td>1864</td>
<td>865,322,642</td>
<td>264,626,771</td>
<td>(600,695,871)</td>
</tr>
<tr>
<td>1863</td>
<td>714,740,725</td>
<td>112,697,291</td>
<td>(602,043,434)</td>
</tr>
<tr>
<td>1862</td>
<td>474,761,819</td>
<td>51,987,456</td>
<td>(422,774,363)</td>
</tr>
<tr>
<td>1861</td>
<td>66,546,645</td>
<td>41,509,931</td>
<td>(25,036,714)</td>
</tr>
</tbody>
</table>

Of course, the largest portion of the rise in expenditures was the combined cost of the War and Navy Departments, which amounted to almost $1.2 billion in 1865 alone.171 Another significant factor, though, was the need to service the mounting debt the Union incurred during the war through the issuance of Treasury notes and other mechanisms. During 1861, in the earliest stages of the war, this amounted to approximately $4 million.

---

169 To illustrate the drop in tariff revenues as a result of both the war and the Panic of 1857, the total customs receipts number was $53,025,794 in 1860 and it was $64,022,863 in 1856. Id.
170 Id. at 300, Series P 99-108.
171 Id.
or six percent of total expenditures. By 1866, however, after the war had concluded, the number has risen to more than $133 million, or more than twenty-five percent of total expenditures.\textsuperscript{172}

IV. Reconstruction

Although the end of the war did not end the significant debt service obligations, there was substantial sentiment to quickly return to the pre-war revenue system. As Randolph Paul observed, the same people who were pleading for both the Union and the Confederacy to substitute taxation for loans during the war were the first to reverse course once the war was over.\textsuperscript{173} Two of the most controversial taxes – the manufacturers tax and the income tax -- were early targets. Representative Frederick Pike of Maine reported in 1866 that they had received “petitions from struggling manufacturers . . . from all quarters of the land, asking for relief.”\textsuperscript{174} The New York Times sided with these manufacturers, arguing that “the [three percent excise tax on all domestically manufactured goods] tax must be abolished at once and completely” on smaller manufacturers and should be phased out over a three-year period on larger manufacturers to “give temporary assistance to the treasury, while our military and naval establishments are gradually being reduced to the peace standard.”\textsuperscript{175}

Similarly, the calls for the repeal of the income tax emerged early on. Arthur’s Home Magazine wrote “[i]t is to be hoped that Congress will, at an early date, entirely repeal the Income Tax, or modify it in many particulars. It was from the very beginning

\textsuperscript{172} Id.
\textsuperscript{173} Paul, supra note 3, at 22-23.
\textsuperscript{174} Cong. Globe, 39\textsuperscript{th} Cong., 1\textsuperscript{st} Sess. 2783 (1866).
\textsuperscript{175} “Taxes on Manufacturers,” N.Y. Times, August 27, 1865, at 4.
a very unpopular tax, and has been rendered doubly so by the publication of incomes, which we have always regarded as a trespass upon private rights of a very grave character.”176 While the calls for income tax reduction were not universal, with merchants and importers viewing the income tax as a potential permanent replacement for the tariff, groups such as bankers and manufacturers kept up the pressure in opposition.177 The calls for reduction were not limited to these two items, however, as many of the objects of internal taxation became sources of irritation when the need for fiscal sacrifice had passed.178

The practical reality was that tax evasion, which many viewed as a self-help form of tax reduction, was only likely to increase with the cessation of hostilities. The whiskey tax was particularly susceptible to evasion, in part because of corrupt or naïve revenue officers.179 A law enforcement source reported that "[t]he best evidence that the Government is being shamefully defrauded by distillers is found in the fact that no whiskey is sold at the price of the tax."180 A similar problem was evident in the other major sin tax on tobacco. According to one report, "[t]he amount of revenue which the

177 Ellis, supra note 138, at 228; Ratner, supra note 40, at 112.
Cotton growers were not the only ones to voice their displeasure. There were protests by shipbuilders, publishers, and manufacturers of wool, among many others. See, e.g., "The Worsted Manufacture and the Tax on Wool," N.Y. Times, January 17, 1867, at 4; "Convention of Shipbuilders in Maine," N.Y. Times, January 31, 1867, at 5; "The Tax on Advertisements," Chicago Tribune, February 3, 1867, reprinted in N.Y. Times, February 16, 1867, at 5.
179 See, e.g., "Whiskey Frauds and Sharp Revenue Officers," N.Y. Times, June 2, 1867, at 4; "Frauds on the Revenue -- New System for Collecting the Tax on Liquors," National Intelligencer, August 8, 1867, reprinted in N.Y. Times, August 9, 1867, at 5 ("The distillers cannot be expected to be more honest then [sic] trusted agents of the Government.").
government is swindled out of, in the cigar and tobacco business, is perhaps beyond belief."181 Evasion had become so common it was a suitable topic for humor. The Saturday Evening Post noted the “singular coincidence” that gold watches had disappeared from record “at the precise moment” when a tax was imposed. The Post sarcastically suggested “[t]he sudden vanishing of so much valuable property should be a matter of public concern, and we desire to directed toward it the attention of all who are interested, in the hope of obtaining some explanation of this remarkable phenomenon.”182

As a consequence of this evasion and the difficulty and costliness of collection, there were many instances of taxes that reportedly did “not pay the expense of collection.”183

Thus, starting in 1866 and continuing through to the end of reconstruction, Congress systematically reduced and refashioned the size of the post-war revenue system. Initially, as the nation attempted to struggle with its post-war debt, the income tax was retained and even broadened slightly,184 but this was under the condition that the tax would only be applied “until and including the year 1870 and no longer.” The general sentiment was that when prices had returned to their normal levels and the burdens on manufacturers lifted, the income tax could be removed as well.185 Under the 1865 Act, political economist David Wells had been appointed as Special Commissioner of the

182 “What Has Become of the Gold Watches?” Saturday Evening Post, June 2, 1866, at 3.
184 See Act of July 13, 1866, 14 Stat. 138 (applying the 10% rate on incomes above $5000 to formerly exempt government officials).
185 Representative Pike, responding to a proposal by Representative Morrill to replace the graduated rate schedule with a flat five percent tax on income over $1000, said “[e]very laboring man pays a tax upon what he eats, drinks, and wears. And until we come to the point of relieving the great body of people in the country from onerous taxes upon everyday’s consumption it is a question whether or not the men who are able to pay should not pay this increased proportion of their income to the General Government.” Cong. Globe, 39th Cong., 1st Sess. 2783 (1866).
Revenue to study a thorough revision of the system. He issued a report in 1866 that recommended a reduction in internal excises taxes and tariffs to address the inflation blamed for post-war economic stagnation. These recommendations were largely adopted over the next several years. As Harry Smith observed, "[n]ow that a reduction in revenue was possible, almost every member of Congress was interested in having some particular tax reduced or repealed because it affected his constituents." In 1866, many internal excises were reduced or repealed altogether, including all taxes on apparel, sugar, advertisements, and cigars. In 1867, Congress repealed many internal excise taxes while also replacing the graduated rate feature on the income tax with a flat five percent rate and raising the exemption to $1,000. Congress repealed the controversial excise tax on cotton in February of 1868 and, a month later in a move considered "sudden and unexpected," repealed all taxes upon manufactured goods, except for those imposed on gas, tobacco, liquor, and banks.

As the date for the expiration of the income tax drew near, an odd coalition formed to grant it a reprieve. While anti-income tax sentiment was as high or higher than it had ever been, income tax supporters aligned with manufacturers, who sought a

---

186 Wells had served on an advisory body to Treasury Secretary Hugh McCulloch prior to his appointment. See Stanley, supra note 1, at 44.
187 Id; Smith, supra note 7, at 235 (describing the recommendations of the special reports of the commission); Howe, supra note 14, at 72. The wholesale price index, which had stood at 89 in 1861, was still at 174 by 1866 after a wartime peak of 193 in 1864. Historical Statistics, supra note 16, at 232, Series L 1-14. More alarmingly, under the Federal Reserve Board of New York’s cost-of-living index, 1866 was at an all-time high of 103, after being at 63 in 1861. It was not until 1912 that the index even approach that high again. Id. at 235, Series L 36-39.
188 See Smith, supra note 7, at 238.
192 See Ratner, supra note 40, at 122; Ellis, supra note 138, at 229.
return to a more explicitly protective tariff. From the perspective of antebellum politics, one might have expected income tax supporters to be anti-protectionist rural dwellers or southern agrarians. What then accounted for the income tax/tariff alliance? One possible explanation is that pro-tariff manufacturers, recognizing the zero-sum game of tax reduction politics, viewed the maintenance of some form of income tax to be preferable to the continued existence of heavy internal duties. One Democrat representative, Stevenson Archer from Maryland, attacked this alliance, suggesting that the income tax would be retained beyond 1870 "for the benefit of manufacturers and high tariff men, who control large bodies of voters, maintaining this burden in order to remove more special taxes from their shoulders." Ultimately, after a battle in the Senate in which proposals for replacing the income tax revenue with taxes on sugar and on gross receipts were rejected, Congress extended the income tax for an additional two years at half of its previous rate and twice the exemption. This provided the revenue to repeal internal duties on items such as carriages, billiard tables, boats, barges, and passports, as well as stamp taxes on mortgages and most promissory notes.

193 See Stanley, supra note 1, at 51.  
194 Indeed, the income tax and tariff were often juxtaposed during commentary on the issue. See Carl Benson, "Things of To-Day: The Income Tax," The Galaxy. A Magazine of Entertaining Reading, IX(2) (February 1870): 267 ("we may conclude that so long as must have quasi war taxes, that is, taxes which are results of the war, the income tax is as fair as any which can be devised. Personally, no doubt, it is a nuisance; individually, every man would rejoice to get rid of it; but when we consider the other burdens pressing on the people, we must acknowledge that some of these demand prior attention. The tariff, for instance, which, besides favoring certain classes, caused a huge waste that does no possible good to any human being.").  
195 On the zero sum game many felt they were operating under, see "The Income Tax," Merchants' Magazine and Commercial Review 62(5) (May 1870): 321 ("The revenues are now sufficient without [the income tax]; and Congress may either repeal it, and retain all the rest of the taxes, or may retain it, at least in part, and so have a surplus to apply in the remission of other duties.").  
197 Act of July 14, 1870, 16 Stat. 257. See Ratner, supra note 40, at 126-27. The income tax rate stood at 2.5 percent on incomes in excess of $2000. The publicity requirement was also repealed.  
This would prove to be only a temporary reprieve for the income tax. While there was some support for continuing the tax, the pressure for additional tax reduction was overwhelming. With internal duties already cut substantially, the income tax was the last target for tax reformers. The New York Times wrote "[w]e have reached the time when the income tax can be no further defended. The people demand its repeal with one voice." Anti-income tax associations in places such as Boston and New York began to explore legal avenues of redress, including efforts to enjoin collectors and to test the constitutionality of the tax. As Representative John Rice admitted, "[t]he question of revenue reform and reduction of taxation is engrossing more of the attention of the people than any other with which this Congress has had to deal." Thus, with the annual budget surplus growing to almost $100 million and prices on non-protected goods beginning their descent from wartime highs, Congress allowed the income to lapse in 1872. Returning full circle to the period before the Civil War, the revenue system now consisted of primarily regressive taxes.

Coupled with the repeal of the inheritance tax in 1871 and the removal of most of the remaining internal excises and stamp taxes in 1872, the end of the income tax marked a fundamental retrenchment of the revenue system. This not only reduced the size of government receipts, but also reduced the size and influence of the government's tax reach. Under the 1872 Act, Congress limited the Office of Internal Revenue to 80
districts, resulting in the abolition of approximately 280 collectors and assessors' offices, a move one contemporary source thought to be "the most important, in an economical point of view, in the entire act." Even tariff rates were reduced significantly, both on protected goods such as cotton, wool, metals, paper, glass, leather, coal and steel, as well as on nonprotected items such as coffee and tea under the guise of affording individuals a "free breakfast table." It was only the Panic of 1873 that allowed protectionists to seize the opportunity to return rates to wartime levels. As a result of the 1870 and 1872 Acts, Federal receipts began a continual downward march that continued to shrink the size of the government over the next decade. It was not until the end of the nineteenth century during the Spanish American War that revenues rose back to 1870 levels.

IV. Civil War Tax Legacy

Given the fundamental retrenchment in the revenue system, it would appear that the Civil War experience acts as a counterexample to the general notion that the war experience results in a ratcheting up of the tax system's influence and, as a consequence, the state's influence in modern society. Most of the Civil War-era taxes were repealed, rather than merely reduced, at the end of Reconstruction. Perhaps the most famous of the new Civil War-era taxes from the perspective of the modern observer—the income tax—

206 Ashley, supra note 6, at 188; Ratner, supra note 40, at 133.
207 In effect, protectionists were not hurt significantly by the 1872 reductions because the largest reductions came from elimination of duties on revenue items. The drop in revenues as a result of the Panic, therefore, meant that increases in rates came in the protected items still subject to tariffs. Ashley, supra note 6, at 188-89.
208 See Historical Statistics, supra note 16, at 297, Series P 89-98. Total receipts, which stood at more than $411 million in 1870, dropped over $150 million by 1878.
209 Between 1899 and 1899, receipts rose from $405 million to just under $516 million. Id. These numbers are not inflation-adjusted.
was discarded with such firmness that it was not adopted again for another quarter century. Rather than leveraging the wartime taxes to fund a larger government, the nation appeared to seek a return to its pre-war status. As Lawrence Friedman observed, "[t]he victorious North was anxious to return to normalcy, at least in money matters. . . . Indeed, the small scale of the federal government did not seem to call for severe and exotic taxes."\textsuperscript{210}

While the Civil War tax program was sharply scaled back during Reconstruction, it was not repealed completely. The Office of Internal Revenue, which had been created as a wartime department of the Treasury, had a major post-war role. Most internal taxes had been repealed in Reconstruction, but substantial excise taxes on liquor and tobacco, remained. These funded nearly ninety percent of internal revenues between 1868 and 1913,\textsuperscript{211} making up almost one-third of all federal revenues.\textsuperscript{212} While the Office was still small, it was growing in size and influence. By 1877, the Office had 188 full-time employees and constituted the second largest civil service operation in the South.\textsuperscript{213} Especially as the economy struggled to recover from the Panic of 1873, the advent of internal revenues proved particularly important in allowing the Federal Government to assume its somewhat heightened role in post-war society.

These surviving features of the wartime tax system helped the State assume a greater role in postwar society. One prominent example is in the area of veterans' pensions. In order to recruit volunteers at the outset of the war, Congress had promised

\footnotesize
\textsuperscript{210} Friedman, \textit{supra} note 4, at 565.
\textsuperscript{211} Id.
\textsuperscript{212} Historical Statistics, \textit{supra} note 16, at 297, Series P 89-98.
\textsuperscript{213} Bensel, \textit{supra} note 47, at 346, 396.
the same veterans pensions available to members of the regular army.\textsuperscript{214} As a result, even before the end of the war, disabled Union soldiers or widows of the deceased began to file for veterans' pensions.\textsuperscript{215} While the Pensions Office had existed prior to the Civil War to administer pensions for veterans of the War of 1812, the costs began to mount sharply during Reconstruction. Veterans' pensions, which had only amounted to approximately $853,000 in 1862, skyrocketed to over $34 million by 1871 or almost as much as the entire amount spent on the War Department for that year.\textsuperscript{216} By 1870, nearly two hundred thousand pensioners, both former soldiers and their widows and dependents, were on the rolls at the Pensions Office. Effectively, it was the first large-scale progressive Federal social welfare program.

A second area of growth in the post-Civil War government was in administering public works projects during Reconstruction and beyond. Throughout most of the latter half of the nineteenth century, federal expenditures on items such as river and harbor improvements and construction of public buildings, armories, and arsenals, amounted to approximately $14 million annually.\textsuperscript{217} In the District of Columbia alone, federal spending on public works projects amounted to $112 per capita by the end of Reconstruction.\textsuperscript{218} Military construction projects, such as a campaign to build naval vessels that picked up in the decade after Reconstruction ended, only added to the total

\textsuperscript{214} By contrast, in the Confederacy President Davis opposed any national pension legislation and even vetoed a bill to create a home for veteran soldiers. Bensel, \textit{supra} note 47, at 166.
\textsuperscript{216} See Historical Statistics, \textit{supra} note 16, at 301, Series P 99-108. The War Department had itself been severely scaled back by this point, operating primarily as a police force for the frontier to battle Indian tribes.
\textsuperscript{218} Bensel, \textit{supra} note 47, at 344 n. 68.
spent on civilian building contracts.\textsuperscript{219} While these public works projects were susceptible to corruption, the alliance of the State and government contractors could not have been possible without the concomitant growth in the federal revenue system.

Not only did the remaining internal taxes fund new government programs, but they themselves acted as new forms of government regulation. Part of the popularity of the liquor and tobacco taxes was that they appeared to sanction activities deemed sinful and unhealthy by a large segment of the populace.\textsuperscript{220} This was one of the grounds for their adoption in 1862. While the general view was taxes on the "vices of men" were "the most reliable of all taxes, as it takes centuries to change or eradicate" them, the understanding was that if they deterred some vice in the process that would not be such a bad thing.\textsuperscript{221} In fact, according to Representative Morrill, the rate on ales and beers were deliberately set lower since "they may be regarded as less unhealthful than spirits."\textsuperscript{222}

More broadly, the Civil War experience laid the groundwork for the subsequent adoption of many of its innovations, including both the inheritance tax and the income tax. During the antebellum era, income taxation was a topic of discussion but never a subject of a serious proposal in Congress. By contrast, between 1873 and 1879, there were fourteen different income tax bills introduced in the House. In 1878, the Greenback movement’s new party aligned with the Labor Reform Party in adopting a platform favoring a graduated income tax.\textsuperscript{223} While none of these bills achieved the two-thirds vote required to overcome the opposition of House leadership, they often were supported

\textsuperscript{219} Id. at 615, 618.
\textsuperscript{221} Cong. Globe, 37th Cong., 2d Sess. 1194-95 (March 12, 1862) (statement of Representative Justin Morrill, D-Vt.) ("The consumption will not be seriously checked; and, if it could be, such a result would bring upon us no national disgrace.").
\textsuperscript{222} Id.
\textsuperscript{223} Id. at 608; Ellis, \textit{supra} note 138, at 231.
by substantial majorities.\footnote{Ellis, \textit{supra} note 138, at 232. Despite the popular sentiment in favor of the income tax in the House, it did not appear that it had any likelihood of passing the Senate. Ironically, one of the leading opponents was David Wells, the former Special Commissioner of Revenues and a member of the 1865 Revenue Commission that had advocated retention of the Civil War income tax. See David Wells, “The Communism of a Discriminating Income Tax,” \textit{North American Review} 130 (March 1880): 236.} Joseph Pulitzer forcefully demanded the return of both the income and inheritance taxes when he took over the \textit{New York World} in 1883.\footnote{Id.} The income tax also had near unanimous support in the academic community among both senior economists such as Amasa and Francis Walker and Arthur Perry, as well as young colleagues such as Richard Ely and Edwin R.A. Seligman.\footnote{The tax was struck down by the Supreme Court in Pollock v. Farmers’ Loan & Trust Co., 157 U.S. 429 (1895).} The Civil War had effectively legitimized both the income tax and the inheritance tax as potential sources of revenue. During the income tax’s brief revival in 1894,\footnote{Act of June 13, 1898, ch. 448, § 27, 29, 30 Stat. 448, 464. See “Taxes to Wage a War,” \textit{Wash. Post}, May 17, 1898, at 4A.} the 1898 adoption of the inheritance tax in the midst of the Spanish-American War,\footnote{Act of October 3, 1913, 38 Stat. 114.} and the enactment of the modern income tax in 1913 after the passage of the Sixteenth Amendment,\footnote{David Wells, “An Income Tax: Is It Desirable?” \textit{Forum} 17 (1893): 1, 3.} the Civil War experience provided valuable support.

Not only did the income tax experience offer legitimacy, but it also provided lessons for the future design of the income tax. During the Civil War, the income tax was much reviled because of its purportedly “arbitrary and inquisitorial method.”\footnote{Kennan, \textit{supra} note 40, at 256.} Due to evasion, it may not have reached more than ten percent of the country’s actual taxable income.\footnote{Ellis, \textit{supra} note 138, at 235.} By contrast, in the areas in which the Union employed stoppage-at-the-source methods, such as the dividends tax and the withholding tax on government salaries, the government avoided the inquisitorial features of the regular income tax and was
considered relatively efficient and successful. In 1894, when the income tax was being considered again, several proponents, including President Grover Cleveland, Treasury Secretary John Carlisle, and House Ways and Means Committee Chair William Wilson, suggested some form of corporate tax similar to the Civil War-era dividends tax. The “increased facility of collection” was one of the principal benefits cited. Eventually, the stoppage-at-the-source method became one of the most important features of the current income tax when it was applied broadly in World War II. The Civil War tax system may have been a product of the war, but it had a lasting legacy even when the patriotic fervor for fiscal sacrifice had dissipated.

---

232 While the income tax implemented through withholding methods was less controversial the personal income tax, the latter did raise more total revenues. See Hill, *supra* note 2, at 450 (noting that while the former raised between $8-$9 million annually, the personal income tax raised $14 million from the 1862 Act alone).


TAXATION & WAR FINANCE DURING THE KOREAN AND VIETNAM CONFLICTS

Kirk J. Stark
Professor of Law
UCLA School of Law

Draft of March 2006

Preliminary and Incomplete
Please do not quote/cite/circulate without author's permission.
TAXATION & WAR FINANCE DURING THE KOREAN AND VIETNAM CONFLICTS

Kirk J. Stark*

Abstract ................................................................................................................................................. 3

Introduction ........................................................................................................................................... 4

I. Korean War & U.S. Tax Policy ......................................................................................................... 7
   A. Background of U.S. Involvement in the Korean Conflict ..................................................... 7
   B. Anticommunist Fervor in Congress: 1949-52 ................................................................... 9
   C. Tax Requests and Legislative Submission: 1950-51 ..................................................... 11

II. Vietnam War & U.S. Tax Policy .................................................................................................... 18
    A. The Vietnam War –Background & Escalating Costs: 1961-1967 ................................ 19
    B. The Vietnam Surcharge Controversy: 1967-68 ........................................................... 21
    C. The Linkage of Surcharge Extension and Tax Reform: 1968-69 ......................... 26

III. Conclusion: Striking While the Iron is Hot .............................................................................. 28

Appendix ............................................................................................................................................... 30

* Professor of Law, UCLA School of Law.
Abstract

This chapter will examine the different political responses of the U.S. government to the war finance problem during the Korean and Vietnam conflicts. Particular attention will be given to the differences in political consensus regarding U.S. involvement in the two wars and how those differences influenced the prospects of war-related tax legislation. Whereas in 1950 there was substantial consensus in both political parties regarding the need for a unified front against communist expansion both at home and abroad, by the mid-1960s that consensus had largely come undone. These changes had significance for the tax policy response to war-related revenue needs. At the outset of the Korean War, President Truman was able to rely upon “crisis politics” and appeals to “universal sacrifice” to ensure swift passage of tax legislation to raise the revenues required for the war. The result was the Revenue Acts of 1950 and 1951 and the Excess Profits Tax of 1950. By the mid-1960s, however, the “crisis” that was the target of the anti-communist movement no longer held the same degree of political sway. Moreover, President Johnson feared that any calls for sacrifice to support the war effort would jeopardize political support for his cherished Great Society programs. Consequently, Johnson delayed seeking increased taxes for Vietnam until he was effectively forced to do so. By that time, however, there was extraordinary reluctance in Congress to adopt (and later extend) the controversial Vietnam War “surcharge.” Recognizing the Administration’s need to fund the war effort, reform-minded lawmakers used their support for the surcharge to force the issue of broader tax reform. Thus, the Tax Reform Act of 1969, a massive piece of tax reform legislation whose effects are with us still today, was an outgrowth of the Administration’s pursuit of an unpopular war. These two episodes in U.S. political/military history reveal how war can serve a catalyst for fiscal change. Ironically, however, it was the (relative) unpopularity of the Vietnam conflict that prompted more fundamental changes in the tax system.
INTRODUCTION

The Korean War (1950-1953) and the Vietnam War (1964-1975) offer unique insights into the special demands placed on taxation and fiscal policy during times of war. At first blush, the two episodes would seem to have much in common in terms of their potential significance for the tax system. Both wars were limited military actions motivated by the U.S. “containment” policy of the Cold War era; neither involved anywhere near the magnitude of fiscal effort required during the Civil War or the two World Wars. Moreover, both the Korean War and the Vietnam conflict were preceded by substantial reductions in the overall tax burden. In the years immediately following the end of World War II, Congress introduced major tax reductions through the enactment of the Revenue Acts of 1945 and 1948. Similarly, the Revenue Acts of 1962 and 1964 substantially reduced the federal tax burden on both individuals and businesses. In both instances, therefore, the increased military expenditures required fresh legislative attention, and reversal of prior years’ tax reductions, to ensure sufficient revenues for the war effort.

Despite these similarities, there are some obvious differences between the two wars that deserve attention. The conflict in Korea enjoyed widespread popular and legislative

---

1 The dates used here are those used by the U.S. Department of Veteran Affairs. See U.S. DEPT. OF VETERAN AFFAIRS, AMERICA’S WARS (November 2004) (www1.va.gov/opa/fact).

2 Revenue Act of 1945 (Ch. 200, 59 Stat. 264; Ch. 210 & 211, 59 Stat 294; Ch. 453, 59 Stat. 556); Revenue Act of 1948 (Ch. 168, 62 Stat. 110). See Figure 2 in the Appendix, reproduced from Jerry Tempalski, Revenue Effects of Major Tax Bills, Office of Tax Analysis Working Paper 81 (July 2003).

3 Revenue Act of 1962 (P.L. 87-834, 76 Stat. 960); Revenue Act of 1964 (P.L. 88-272, 78 Stat. 19). See Figure 2 in the Appendix, which shows revenue estimates for all tax legislation from 1940-1967.
support—at least at the outset. The North Korean invasion of South Korea on June 25, 1950, came at the height of McCarthyism, during a time when both political parties were eager to earn distinction as fervent anticommunists. By contrast, the conflict in Vietnam, which began largely under the radar screen, would eventually become the most unpopular military action in American history and took place during an era of extraordinary political dissensus. Thus, a careful comparison of the two periods offers the opportunity for a natural experiment of sorts—a chance to view the problem of war finance in the competing contexts of (relative) harmony and discord.

Not surprisingly, the extent of political disagreement regarding the wisdom of military action exerts a fundamental influence on the legislative process: “war taxes” are less likely to garner political support when the war itself is so hotly contested. Thus, an administration pursuing an unpopular military action may be required to offer legislative concessions, in the form of changes to the tax system, in order to secure funding for the war effort. By contrast, political consensus over the wisdom of the war allows revenue measures to proceed with less horse-trading. While the limited historical evidence presented here should be interpreted with caution, an examination of these two periods suggests that in the case of controversy and discord, the resulting fiscal measures may prove to be more durable in the post-war environment. The intuition behind this claim is that while the taxes used to finance an unpopular military action may themselves be temporary, the legislative concessions required to secure the enactment of those taxes

---

4 Indeed, the term “McCarthyism” originated just four months before the North Korean invasion in a February 1950 Washington Post editorial cartoon by Herbert Block.
may outlive the conflict. Ironically, therefore, tax law changes arising from an unpopular war may be more likely to stick with us over the long term.

As will be explained in further detail below, this is essentially the story of the Korean and Vietnam conflicts. Following the U.S. decision to enter the conflict in Korea, Congress acted quickly in response to President Truman’s request for new taxes to finance the effort. In less than six months, Congress responded with two major tax bills—the Revenue Act of 1950 and the Excess Profits Act of 1950. Fifteen years later, however, President Johnson faced an entirely different political dynamic. Initially reluctant to seek new taxes from Congress, Johnson was eventually forced to request a income tax “surcharge” to help pay for the escalating war effort. By the time Johnson acted, however, Congress was bitterly divided over the wisdom of U.S. actions in Southeast Asia. The delay weakened the President’s hand, giving reluctant legislators the leverage to extract concessions to allow the surcharge to go forward. When Nixon later requested an extension of the surcharge, Congress demanded fundamental tax reform in exchange for its approval. The product of that legislative compromise was the Tax Reform Act of 1969 and a host of complex provisions that are still with us today—including the much-maligned alternative minimum tax.

\footnote{5 Discussion infra at Part I.C.}
I. KOREAN WAR & U.S. TAX POLICY

A. Background of U.S. Involvement in the Korean Conflict

U.S. involvement in the Korean conflict had its origins in World War II and the Japanese occupation of the Korean peninsula from 1910-1945. Two days after the bombing of Hiroshima on August 6, 1945, Soviet troops entered the Korean peninsula from the North. In September, U.S. troops entered the peninsula from the South. On September 6, 1945, the U.S. and U.S.S.R agreed that the country should be divided along the 38th parallel. Although there were efforts in the ensuing three years to unify the peninsula, by mid-1948, it had become apparent that this would not occur. In August 1948, the Republic of Korea (South) was established. In September of that year, the People’s Republic of Korea (North) was established. Roughly eighteen months later, on June 25, 1950, North Korea launched an invasion of the South at several points along the border.6

The Korean War lasted from June 1950 to July 1953. The U.S. military effort, undertaken officially through the United Nations, began with a series of strategic setbacks in July and August 1950.7 The tide of the war took a dramatic turn, however, with General Douglas MacArthur’s surprise amphibious landing at Inchon on September

7 Id. at 215-216.
15, 1950. MacArthur was, of course, the leading figure—and the most controversial figure—of the Korean conflict. By September 26, 1950, UN forces had retaken Seoul and were poised to destroy the North Korean army. Buoyed by these successes, the Truman administration decided that the next step should be the reunification of the Korean peninsula. With the official backing of the United Nations, U.S. troops crossed the 38th parallel and began repelling North Korean forces toward the Yalu River in the north. Given these developments, there was good reason for optimism. At a special meeting at Wake Island in mid-October, 1950, General MacArthur assured Truman that the war would be over by November.9

This was the point at which General MacArthur began taking actions beyond the scope of the mandate from Washington. As Patterson explains, in late October MacArthur disobeyed orders from Truman and ordered American troops close to the Yalu River along the North Korean/Chinese border. This turned out to be a significant miscalculation. In late November, Chinese forces began pouring into North Korea. The war had entered a new stage, with direct Chinese involvement posing a significant military challenge to the UN forces. Military losses were significant. For example, during the last three days of November 1950, the U.S. 7th Division suffered 5,000 casualties—one-third of its total force.10 By mid-December Chinese forces had crossed

---

8 Id. at 219.
9 PATTERSON at 220.
10 PATTERSON at 222.
the 38th parallel and announced their intention of unifying the peninsula under communist control.\textsuperscript{11}

With these setbacks, the US was forced to abandon its strategy of unifying the peninsula. By mid-March 1951, the Truman Administration had decided to pursue peace negotiations with the Chinese. MacArthur balked, issuing his own statement indicating his willingness to meet with the Chinese, but also emphasizing that he was prepared to invade China directly with U.S. troops. MacArthur was removed from his post on April 11, 1951 and peace talks began in July that year. The fighting lasted for another two years until late July 1953, when the Chinese and North Koreans agree to repatriate all UN prisoners of war.\textsuperscript{12} The Korean War Armistice Agreement, dated July 27, 1953, provided for a cease-fire, withdrawal of troops and establishment of the demilitarized zone along the 38th parallel. By the end of the war, 33,629 American soldiers had been killed and 103,284 wounded.\textsuperscript{13}

\textbf{B. Anticommunist Fervor in Congress: 1949-52}

Popular and political support for the military action in Korea ebbed and flowed over the course of the war with the rollercoaster of events on the peninsula. During those first few weeks in the middle of the summer in 1950, however, political support for the Administration’s efforts was at its zenith. Despite his decision not to consult Congress

\begin{thebibliography}{9}
\bibitem{patterson} PATTERSON at 234.
\bibitem{patterson2} PATTERSON at 208.
\end{thebibliography}
before committing U.S. troops, Truman’s actions garnered strong support on both sides of the aisle. These political dynamics naturally set the stage for the financing of the war.

As Witte explains in his comprehensive history of US tax politics, “only in those early months of almost hysterical reaction to crisis can revenues easily be raised in the United States.” This observation holds especially true for the months following the invasion of South Korea in late June 1950. It is important to keep in mind the domestic political context surrounding the onset of the Korean conflict. The summer of 1950 was the height of McCarthyism, a time during which anticommmunist fervor pervaded American politics. Less than a year before the North Korean invasion, China had “fallen” to the communists with Mao Zedong’s proclamation of the People’s Republic of China on October 1, 1949. Republican critics of the administration (joined by some Democrats, including John Kennedy, then a young member of the House of Representatives from Massachusetts’ 11th congressional district) charged Truman with having “lost China.” During the first half of 1950, headlines featured a number of figures that today we associate with the “Red Scare” of the McCarthy era—Alger Hiss, Klaus Fuchs, Julius and Ethel Rosenberg. On February 9, 1950, Senator Joe McCarthy made his infamous Wheeling speech in which he claimed that there were 200 communists

16 In January 1950, Alger Hiss was convicted of perjury and sentenced to five years in federal prison. Less than two weeks later, Klaus Fuchs confessed to serving as a Soviet spy, passing on sensitive information relating to the Manhattan Project. It was the Fuchs investigation that lead to the Rosenbergs.
working at the U.S. State Department. In short, at the onset of hostilities in Korea, there was extraordinary political unity regarding the need to fight global communism.

C. Tax Requests and Legislative Submission: 1950-51

In this context of rabid anticommunism, it is not surprising that the President would be able to garner strong support to raise taxes for a war to contain the communists. In the weeks following the invasion of South Korea, President Truman and the two Congressional tax-writing committees acted quickly to ensure adequate tax financing for the war effort. These decisive measures reflected a political consensus that the war effort required “universal sacrifice” – a point driven home by John Foster Dulles in a New York Times article in July 1950. Having visited Korea as a special representative of the President Truman in June, Dulles wrote that “The time for sacrifice and discipline is here. In fact it has been here for long, but it took the attack on Korea to open our eyes. Many now will risk their lives before the hard battle of Korea is won. Most of us will have to work longer hours and with more intensity. We shall all have to give up some material enjoyments and be more frugal in our living. There will be fewer automobiles, television sets and gadgets to buy and there will be bigger tax bills to pay.”

Revenue Act of 1950. The first step toward the “sacrifice and discipline” of which Dulles wrote was the Revenue Act of 1950. As can be seen in the Figure in the Appendix, this legislation resulted in the largest tax increase since the Revenue Act of

17 John Foster Dulles, ‘To Save Humanity From the Deep Abyss’. NEW YORK TIMES, SM3 (July 30, 1950).
1942. Originally introduced as H.R. 8920 in mid-June 1950, the Revenue Act of 1950 had originated as an excise tax reduction bill. With the onset of hostilities on the Korean peninsula, however, legislators converted the measure to substantially increase revenues.\(^{18}\) As the Senate Report noted,

“The military action in Korea coupled with substantial increases in defense and related expenditures has made it necessary to convert the excise tax reduction bill passed by the House in June of this year into a bill to raise revenues. The bill as amended by your committee will increase tax liabilities by 4.5 billion dollars a year when fully effective, and will increase collections in the fiscal year 1951 by about 3 billion dollars. It is not anticipated that these increases will be of sufficient size to offset the new defense and related expenditures. However, this bill accomplishes all that can be done quickly.”\(^{19}\)

While the final legislation incorporated certain of the revenue-raising provisions contained in the original pre-war bill, the Senate Report explains that the major source of revenue from the 1950 Act would come from higher corporate and individual tax rates.\(^{20}\) The corporate tax rate was increased from 38 percent to 45 percent, while individual rates were changed to reverse the reductions of the 1945 and 1948 legislation. As a result of the Revenue Act of 1950, individual rates would range from a low of 20 percent to a high of 91 percent.\(^{21}\) The increased rates would apply to approximately one-half of 1950

---


\(^{20}\) Id.

\(^{21}\) Id.
income for corporations and one-quarter of 1950 income for individuals.\textsuperscript{22} There was no effort to dance around the fact that this was a significant and immediate increase in taxes for most Americans. The Senate Report stated the effect bluntly: “These changes in the corporate and individual income-tax rates involve few technical problems and there is general agreement that these rates must be raised in view of the new expenditures required by the crisis in international affairs.”\textsuperscript{23}

While passage of the Revenue Act of 1950 was swift and decisive, it was not sufficient to satisfy the full funding requirements of the war effort. According to then Secretary of the Treasury John Snyder, the legislation “was not quite up to what it should have been, but it at least made a step towards meeting the Korean expenditures.”\textsuperscript{24} President Truman signed the Revenue Act of 1950 into law on September 23, 1950, less than 45 days from the Administration’s initial request in August 1950.\textsuperscript{25} Incorporated within the new law was a provision that required Congress to report out an excess profits tax as soon as possible.\textsuperscript{26}

If there was a high point in the fiscal reaction to the Korean conflict, it was at this point in late September 1950. The “rally around the flag” effect was in full force.

\textsuperscript{22} Id.
\textsuperscript{23} Id.
\textsuperscript{24} Oral History Interview with John W. Snyder, Secretary of the Treasury in the Truman Administration, 1946-1953 (July 2, 1969) (available in the Truman Presidential Museum & Library).
\textsuperscript{25} Id. See also WITTE at 138.
\textsuperscript{26} Witte describes this provision as one that “bound the revenue committees of Congress to report a bill establishing a retroactive excess profits tax at the earliest possible opportunity in the next Congress.” WITTE at 138.
MacArthur’s masterful amphibious landing at Inchon on September 15 and victories north of the 38th parallel in the ensuing weeks helped to convert tax politics at home into a political competition over who could most support the war effort. In the final two months of 1950, however, Truman would face his most difficult time in office.27

Excess Profits Tax Act of 1950. Later that year, acting as required by the Revenue Act of 1950, Congress began consideration of an excess profits tax to provide additional funding for the war. While work on this legislation began almost immediately after passage of the Revenue Act in late September, it is important to note that the six weeks immediately following adoption of the Revenue Act were also the six weeks immediately preceding the midterm election of November 1950. This was a crucial election for the Republicans. Democrats retained a majority in both chambers, but the Republicans gained a net five seats in the Senate and twenty-eight seats in the House.28 Significantly, among those defeated was Illinois Democrat Scott Lucas, Senate Majority Leader. This was the election that brought freshman California Senator Richard M. Nixon into the national spotlight and also opened up a path for Texas Senator Lyndon Johnson that would eventually lead him to his position as Senate Majority Leader.29

27 PATTERSON at 222 (citing McCullough’s observation that November and December 1950 were “a dreadful passage for Truman . . . the most difficult period of his Presidency.”).

28 Democrats had held a 12 seat majority in the Senate and a 92 seat majority in the House. Following the November elections, Democrats held a 2 seat majority in the Senate and a 35 seat majority in the House. PATTERSON at 223.

29 Johnson’s good fortune in November 1950 was the defeat of Millard Tydings in the Maryland Senate race. Tydings’ defeat enabled Johnson to assume the Chair of the Subcommittee on Korean War Preparedness, which position gave him a platform for assuming national prominence. Johnson would later serve as Assistant Majority Leader from 1951 to 1953 and Minority Leader from 1953 to 1955. He would become Senate Majority Leader in January 1955, transforming the power and function of that position. For
Following the November election, Truman officially requested excess profits legislation that would generate at least $4 billion per year. In a special lame-duck session, Congress acted immediately by enacting an Excess Profits Tax similar to the World War II tax that had been repealed in 1946. The Senate Report described the legislation as “the second step in the financing of the vastly expanded military program resulting from hostilities in Korea and the critical international situation.” The Act was expected to raise $3.2 billion in 1950 and additional $4-$5 billion in calendar year 1951. As with previous conflicts, justification for an excess profits tax was grounded in the view that business profiting from the war effort should be required to pay a larger share of the war’s costs: “One of the main advantages of an excess-profits tax in periods of large military expansion is that it selects for additional tax those corporations whose profits are higher than they probably would have been in the absence of hostilities and a large military budget.”

Once again, bipartisan enthusiasm for the war effort produced a strong political consensus in favor of “universal sacrifice” through tax increases. Republican Senator Robert Taft promoted Truman’s tax increases for the war. As the New York Times reported, “[d]iscussing what he called the inevitable sacrifices that lay immediately ahead

---

a fascinating account of this period in Johnson’s career, see Robert A. Caro, The Years of Lyndon Johnson: Master of the Senate (2002).

30 Witte at 138.
33 Id.
for Americans, Senator Taft said that in connection with tax increases, everybody would have to cut his standard of living by perhaps 10 per cent.\textsuperscript{34}

\textit{Revenue Act of 1951.} The final piece of major tax legislation enacted during the Korean War was the Revenue Act of 1951. If the Revenue Act of 1950 was the strongest evidence of the “rally around the flag” effect in the postwar era, the 1951 legislation was a preview of the contentious politics that would dominate U.S. war tax debates in the second half of the 1960’s. Whereas the 1950 legislation was introduced and signed into law within less than 45 days, the 1951 legislation was preceded by a full nine months of congressional testimony and considerable horse-trading.\textsuperscript{35}

[Add discussion of 1951 legislation; key provisions].\textsuperscript{36} Battles in the House & Sam Rayburn’s highly unusual speech on the House floor: “\textit{I think the boys in Korea would appreciate it more if we in this country were to pay our own way instead of leaving it for them to pay when they get back.}”\textsuperscript{37} Opposition to additional taxes in Senate by November 1951.\textsuperscript{38}

* * *

\textsuperscript{34} William S. White, \textit{Taft Offers to Aid Truman in Shaping U.S. Foreign Policy}, NEW YORK TIMES (January 10, 1951).
\textsuperscript{37} John D. Morris, \textit{Vote is 185 to 160: Earlier Defeat Reversed as 26 Heed Appeal by White House, Rayburn Voices Warning}, NEW YORK TIMES 1, (October 20, 1951).
\textsuperscript{38} \textit{Taft Renews Call for Korean Truce}, NEW YORK TIMES 62 (November 18, 1951) (“Mr. Taft predicted that the President would have great difficulty in getting any additional tax increases out of Congress.”).
In combination, the Revenue Act of 1950, the Excess Profits Tax of 1950 and the Revenue Act of 1951 increased FY1951 tax revenues by $14.7 billion, an amount which, according to Treasury Secretary Snyder, “materially assisted the Treasury in its program of trying to keep on top of our fiscal and monetary problems without going into deficit financing.”\textsuperscript{39} In retrospect, the U.S. experience in financing the Korean War stands out as the high point of bipartisanship in the postwar era. Indeed, it is hard to imagine any other moment in the past half-century when the U.S. political establishment has spoken with such a unified voice with regard to tax policy.

In a recent study of the economic effects of war finance, economist Lee Ohanian emphasizes the uniqueness of the Korean War experience. Whereas financing of every major preceding military engagement (including the Revolutionary War, the War of 1812, the Civil War, World War I and World War II) depended heavily on government borrowing, expenditures for the Korean War “were financed almost exclusively by higher capital and labor income taxes.”\textsuperscript{40} Viewed in the political context described above, it is not hard to see why Congress acted so promptly in response to Truman’s requests for new war taxes. The political circumstances of 1950-1951 were highly unusual. McCarthyism and the “Red Scare” defined legislative politics during the first six months following the North Korean invasion. Truman exploited the politics of the moment, winning substantial tax increases at the very outset of the war.

\textsuperscript{39} Id.

Truman’s decisive action to seek tax increases for the Korean War effort in the second half of 1950 should come as no surprise. After all, he had earned his national political reputation during his leadership of the famous “Truman Committee” on war preparedness in the early 1940s. Presumably the lessons he learned during his chairmanship of that committee informed his later actions as a wartime President. A similar committee was established in late 1950 to investigate military preparedness for the Korean War. Ironically, however, that committee’s chairman—Texas Senator Lyndon Baines Johnson—would choose a wholly different path as a wartime President fifteen years later. Johnson’s refusal to follow Truman’s example in seeking early and substantial tax financing for the Vietnam conflict set the stage for a more contentious battle over war financing in the late 1960s.

II. VIETNAM WAR & U.S. TAX POLICY

As noted above, there are some important similarities between the fiscal situation in the years leading up to Korean War and the budgetary climate of the first half of the 1960s. Both wars were preceded by significant tax reductions and both would eventually require substantial tax increases to meet the war funding needs. Unlike Truman, however, President Johnson steadfastly refused to seek a tax increase to pay for the war in Vietnam until he was effectively forced to do so in early 1967. By that time, however, opposition to the war had mobilized and Johnson faced an uphill battle in securing a tax increase in a Congress increasingly divided over the war.
A. The Vietnam War –Background & Escalating Costs: 1961-1967

U.S. military involvement in Vietnam began in the early 1960s with the signing of a military and economic aid treaty with South Vietnam and the arrival of U.S. military advisors to the regime of Ngo Dinh Diem. 41 Over the thirty-four months of the Kennedy presidency, the U.S. increased the number of military advisors in South Vietnam to more than 16,000. Yet the major push for U.S. involvement would follow the assassinations of Diem on November 2, 1963 and Kennedy three weeks later.

In the two years following Lyndon Johnson’s ascension to the Presidency on November 22, 1963, U.S. involvement in Vietnam escalated rapidly. 42 At the end of 1963, the number of military personnel in Vietnam was 17,000. That figure increased to 23,000 by the end of 1964, 184,000 by the end of 1965, 450,000 by the end of 1966, and more than 500,000 by early 1968. 43 Obviously a build-up of this magnitude would require substantial resources, yet Johnson was extremely reluctant to seek increased taxes, apparently because he feared losing political support for his Great Society programs. From a budgetary perspective, the obvious question was: could the


43 PATTERTSON at 595.
government finance the Great Society programs and the war in Vietnam without increasing taxes?44

Publicly at least, Johnson was confident that the country could have both “guns” and “butter” without increasing taxes. In his State of the Union address on January 1966, Johnson proclaimed that “[t]his Nation is mighty enough, its society is healthy enough, its people are strong enough, to pursue our goals in the rest of the world while still building a Great Society here at home… Time may require further sacrifices. And if it does, then we will make them. But we will not heed those who wring it from the hopes of the unfortunate here in a land of plenty. I believe that we can continue the Great Society while we fight in Vietnam.”45

It is not clear how forthright Johnson was being with the country in his 1966 State of the Union address. As he later explained in an interview with Doris Kearns Goodwin,

“I knew from the start that I was bound to be crucified either way I moved. If I left the woman I really loved—the Great Society—in order to get involved with that bitch of a war on the other side of the world, then I would lose everything at home. All my programs. All my hopes to feed the hungry and the homeless. All my dreams.”46

44 [note parallels to current situation]. See Alan Auerbach, American Fiscal Policy in the Post-War Era: An Interpretive History 3 (March 2005) (“During the George W. Bush administration, spending in all three areas has grown as a share of GDP, for the first time since Johnson administration’s simultaneous pursuit of the Great Society and the Vietnam War.”) See also Stephen Slivinski, Bush Beats Johnson: Comparing the Presidents, CATO INSTITUTE TAX & BUDGET BULLETIN (October 2005) (noting that President George W. Bush is “an even bigger spender than Lyndon B. Johnson in terms of discretionary spending). See also Julian E. Zelizer, Taxing the Homefront: Congress and Taxation in 1968 and 2003 (describing parallels between tax politics during conflicts in Vietnam and Iraq).
45 LBJ, Annual Message to the Congress on the State of the Union (January 12, 1966).
46 PATTERSON at XX (citing Doris Kearns Goodwin, to whom Johnson made these remarks).
Here Johnson seems to be indicating that it was just a matter of time before he would have to face a trade-off in the question of guns versus butter.

Whatever his actual beliefs, Johnson initially predicted fairly modest expenditures in Vietnam. In his 1966 address, Johnson estimated that fiscal 1967 spending on the Vietnam effort would not exceed $10.5 billion. In addition, he sought $3.2 billion in additional funding for domestic programs.47 Within a year, however, it became apparent that the administration had seriously miscalculated the funding required for the war effort.48 In early 1967, Johnson sought an extra $12 billion in fiscal 1967 appropriations for Vietnam and also estimated that fiscal 1968 spending would be approximately $22 billion.49 Clearly something had to give.

B. The Vietnam Surcharge Controversy: 1967-68

By early 1967, it had become clear that existing revenues would not be sufficient for the mounting expenditures in Vietnam, along with the President’s Great Society programs. Johnson had little interest in tax policy and even less in raising taxes. Yet the inevitability of doing so was increasingly apparent. In his State of the Union Address on January 10, 1967, Johnson proposed a income tax “surcharge” to help fund the war in Vietnam:

47 CONGRESSIONAL QUARTERLY WEEKLY REPORT at 1969 (July 26, 1968).
48 Id.
49 Id.
“I recommend to the Congress a surcharge of 6 percent on both corporate and individual income taxes-to last for 2 years or for so long as the unusual expenditures associated with our efforts in Vietnam continue. I will promptly recommend an earlier termination date if a reduction in these expenditures permits it. This surcharge will raise revenues by some $4.5 billion in the first year. For example, a person whose tax payment, the tax he owes, is $1,000, will pay, under this proposal, an extra $60 over the 12-month period, or $5 a month. The overwhelming majority of Americans who pay taxes today are below that figure and they will pay substantially less than $5 a month. Married couples with two children, with incomes up to $5,000 per year, will be exempt from this tax-as will single people with an income of up to $1,900 a year.”

Johnson seemed to be at pains to emphasize what a minimal sacrifice was involved in his surcharge proposal. He added that even with the tax increase Americans would be paying less than they had been when he took office in late 1963. The Revenue Act of 1964 had lowered taxes significantly (see Figure in Appendix), so much so in fact that, as Johnson put it, paying the proposed 6 percent surcharge meant giving up “slightly more than one-fourth of that tax cut each year in order to try to hold our budget deficit in fiscal 1968 within prudent limits and to give our country and to give our fighting men the help they need in this hour of trial.”

Despite his efforts to portray the surcharge proposal as necessary and reasonable, Johnson faced extraordinary difficulty in getting it enacted. Indeed, this “Master of the Senate” would not see his war financing proposal enacted until a full eighteen months

50 LBJ, Annual Message to the Congress on the State of the Union (January 10, 1967).
51 Id.
later, when he signed into law P.L. 90-364, the Revenue and Expenditure Control Act of 1968, on June 28, 1968. To be sure, there are complicated reasons for the legislative difficulty that Johnson faced in getting the surcharge enacted. To begin, what began as a proposal for a 6 percent surcharge became, by the end of the summer of 1967, a proposal for a 10 percent surcharge. That alone might have been enough to give rise to skepticism about how honest the Administration was being with regard to the costs of prosecuting the war.

More significantly, however, the period from August 1967 to June 1968 proved eventful for reasons going far beyond the controversy over Johnson’s proposed surcharge. Race riots broke out in several U.S. cities, with those in Newark and Detroit being the most severe. Meanwhile, in Vietnam, the number of U.S. soldiers killed, wounded or missing in action jumped dramatically from 1966, yet the outcome of the conflict was less certain than ever. In November 1967, Secretary of Defense Robert McNamara resigned, having expressed to the President increasing doubt regarding the prospects for the U.S. war effort.

The New Year brought darker developments. On January 30, 1968, the Tet Offensive began as Vietcong forces launched several coordinated operational assaults throughout South Vietnam. Then, in a move that shocked the nation, Lyndon Johnson announced on March 31, 1968, that he would no longer seek renomination for the Presidency. Five days later, on April 4, 1968, civil rights leader Martin Luther King, Jr. was assassinated.

---

52 Id.
in Memphis, Tennessee, which lead to further urban riots throughout the country. Two months after that, Robert Kennedy was assassinated in Los Angeles, California.

Given this context, it is no surprise that Lyndon Johnson struggled to advance his legislative proposal to fund the war via a 10 percent surcharge. Johnson’s approval ratings reached all-time lows at precisely the time Congress was taking up the surcharge proposal. Legislators seized the opportunity to extract concessions from the President in the form of domestic spending cuts. In particular, Rep. Wilbur Mills, chair of the House Ways and Means Committee, refused to act on the surcharge unless and until the Administration agreed to substantial reductions in Great Society programs. In the final deal, passage of the surcharge was exchanged for “an appropriations package that reduced spending by a total of $18 billion (out of a $180.1 billion budget).”54 [further discussion here of the clash between Johnson and Wilbur Mills over the Vietnam surcharge].55 Because of the delay in legislative action, the annual budget deficit increased fourfold from FY1966 to FY1967 (from $3.1 billion to $12.6 billion), then more than doubled for FY1968 (to $27.7 billion).56 These fiscal policies contributed directly to the “Great Inflation” of 1965-1984, which has recently been described as “the climactic monetary event of the last part of the 20th century.”57

53 [CQ cite for revised proposal; spending projections].
54 Zelizer, Taxing the Homefront.
57 See Allan H. Meltzer, Origins of the Great Inflation, FEDERAL RESERVE BANK OF ST. LOUIS REVIEW
Johnson signed the Revenue and Expenditure Control Act of 1968 (enacting the surcharge) into law on June 28, 1968. The importance of the 1968 legislation to the U.S. budget situation in the late 1960s should not be underestimated. According to a recent Treasury Department analysis, the Revenue and Expenditure Control Act of 1968 was the largest single-year tax increase in U.S. history since the end of World War II, raising $55 billion in constant 1992 dollars.58

The 1968 law was limited to one year and therefore set to expire at the end of June 1969. Significantly, however, the law contained a provision, introduced in the Senate as an amendment to the bill by New York Senator Jacob K. Javits, that required the Administration to submit to Congress by December 31, 1968, proposals for tax reform. The Treasury Department, under the leadership of Assistant Secretary for Tax Policy Stanley S. Surrey, had been working on tax reform proposals for two years, but with little interest or leadership from an increasingly beleaguered White House. The Javits amendment created an opportunity for tax reform to find a more secure footing on the legislative agenda. Later in 1968, however, Johnson decided not to forward the proposals.59 An agreement was reached that the Treasury proposals would be released at the beginning of the New Year. Tax reform, if any, would have to await the new administration of Richard Nixon, who was scheduled to take office in January 1969.

(March/April 2005).


59 WITTE at 166.
C. The Linkage of Surcharge Extension and Tax Reform: 1968-69

What happened next is a chapter in U.S. tax history that has been repeated many times in recent years. As Witte explains in his political history of the U.S. income tax, “[p]rior to the release of the documents in January 1969, the outgoing Secretary of the Treasury, Joseph Barr, made a highly publicized speech warning of a “tax revolt” based on the inequities in the tax code, particularly the “loopholes” that permitted the very rich to avoid taxation. Employing a strategy used successfully by the Progressives forty years earlier, he listed some of the very wealthy individuals who paid no taxes at all.”60 Modern day tax scholars know this story well—the Barr revelations created a groundswell of political support for tax reform that ultimately lead to the incorporation of the alternative minimum tax in the Tax Reform Act of 1969. What is less well known is the precise connection between the legislative consideration of tax reform and the Vietnam surcharge.

During the first six months of the Nixon administration, pressure mounted to continue the surcharge that was scheduled to expire at the end of June 1969. Administration officials cited inflation concerns and continued spending needs in Vietnam as support for extending the measure at least another year. With both the House and Senate controlled by Democrats, however, Nixon faced an uphill battle on the Hill. On May 14, fifty-four Democrats in the House of Representatives asked Ways and Means Chairman to “delay extension of the 10-percent surcharge until there is assurance of passage of a major tax-

---

60 Id.
reform bill.\textsuperscript{61} In addition, the executive council of the AFL-CIO issued a statement on the same day indicating that it would not support surcharge extension “until it is combined with immediate, substantial and equitable reform of the federal income tax structure.”\textsuperscript{62} It appeared that, unless the Administration was willing to join forces with Democratic liberals in pursuing broad-based tax reform, the surcharge would expire at the end of June as scheduled. On June 12, 1969, the Nixon Administration reached an agreement with House leaders trading off tax reform for an extension of the surcharge.\textsuperscript{63} That was not, however, the end of the story. [describe political back-and-forth between Nixon, Mansfield and other Senate Democrats re: linkage of surtax extension and tax reform]\textsuperscript{64}

The Tax Reform Act of 1969, signed into law by President Nixon on December 30, 1969, incorporated several major structural reforms to the income tax, including repeal of

\textsuperscript{61} CONGRESSIONAL QUARTERLY WEEKLY, May 16, 1969 at page 736 (noting that signers of the letter were all members of the liberal “Democratic Study Group.”).

\textsuperscript{62} Id.

\textsuperscript{63} Administration, House Leaders Compromise on Surtax, CONGRESSIONAL QUARTERLY WEEKLY REPORT 1013 (June 13, 1969).

\textsuperscript{64} Surtax Extension, CONGRESSIONAL QUARTERLY WEEKLY REPORT 1290 (July 18, 1969) (“Mansfield in a July 14 statement on the Senate floor said the surtax bill would probably be defeated on the Senate floor if major tax reforms were not tied to it.”); Oil Depletion, Bank and Utility Tax Advantages Cut, CONGRESSIONAL QUARTERLY WEEKLY REPORT 1315 (July 25, 1969) (“Maj. Leader Mike Mansfield (D. Mont.) had said he would not schedule a vote on extension of the surtax until the House tax reform bill was ready for Senate floor action.”); Surtax Extension, CONGRESSIONAL QUARTERLY WEEKLY REPORT 1316 (July 25, 1969) (Mansfield “remained firm in his resolve to tie tax reform to the surtax even after a July 22 meeting with President Nixon who had been equally firm in his desire to get the surtax extended.”); House Votes Overwhelmingly for Landmark Tax Reform, CONGRESSIONAL QUARTERLY WEEKLY REPORT 1424, 1428 (August 8, 1969) (“Committee added on to the tax reform bill the rest of the surtax extension package.”); House Accepts Senate Compromise on Surtax Extension, CONGRESSIONAL QUARTERLY WEEKLY REPORT 1013 (August 8, 1969).
the investment tax credit, the new alternative minimum tax, a revised income-averaging provision (originally enacted in 1964 and eventually repealed in 1986), new rules for capital gains taxes.

III. CONCLUSION: STRIKING WHILE THE IRON IS HOT

If there is a lesson to be learned from examining U.S. tax policy and war financing during the Korean and Vietnam conflicts, it is that political leaders must “strike while the iron is hot” if there is any hope of enacting tax increases to fund military efforts abroad. President Truman displayed the advantages of this approach in the six months immediately following the North Korean invasion of South Korea on June 25, 1950, acting quickly and decisively to raise capital and labor income taxes. By contrast, President Johnson refused to follow this approach.

The history of Johnson’s struggle to get the Vietnam surcharge enacted between January 1967 and June 1968 may be usefully contrasted with Truman’s request for tax-financed appropriations for the Korean conflict in the late summer of 1950. It is tempting to explain the difference in outcome with the observation that Korea was a popular war while Vietnam was unpopular. But this simplistic view overlooks the fact that the conflict in Vietnam did not start out as an unpopular war. As Historian James Patterson notes, “[u]ntil 1968 Johnson enjoyed considerable bipartisan support in

65 Miscalculations Almost Killed Tax Surcharge, CONGRESSIONAL QUARTERLY WEEKLY REPORT (July 26, 1968).

Congress for his support of the war.” Of course, enjoying bipartisan support for the war does not necessarily mean enjoying bipartisan support for tax increases to pay for the war, as evidenced by the eighteen months it took to get the Vietnam surcharge enacted. Still, one cannot help but think that there were earlier opportunities—perhaps, for example, in the weeks following the Vietcong attack on an American air base in Pleiku in February 1965—that might have given Johnson the chance to act with decisive fiscal leadership in the manner that Truman had in the summer of 1950.

---

67 PATTERSON at 608 (citing Herring, People Quite Apart: Americans, South Vietnamese and the War in Vietnam, 9 DIPLOMATIC HISTORY 14, 39-40 (1990); also Kearns, Lyndon Johnson, 324-327).
Appendix

Figure 2 - Tax Bill Revenue Estimates in Constant Dollars, 1940 - 1967
(Full year effect unless otherwise noted)

* indicates revenue estimate is for first year after enactment

Taxation & Korean War in New York Times (1950)

Fiscal Unpreparedness. NEW YORK TIMES (July 5, 1950).

No Time for Tax Action. NEW YORK TIMES (July 12, 1950).

Realism On Taxes. NEW YORK TIMES (July 14, 1950).

The Front Line of Taxation. The NEW YORK TIMES (December 4, 1950).

Anthony Leviero, President is Blunt. NEW YORK TIMES, A1, December 16, 1950.

Text of President’s Address on ‘Great Danger’ Facing the Nation. NEW YORK TIMES, A4, December 16, 1950.


Second Defense Tax Bill. NEW YORK TIMES (December 27, 1950).


John D. Morris, President Pleads for Tax-Bill Speed. NEW YORK TIMES, A18 (August 31, 1950).

John D. Morris, Decision Reversed on Tax-Rise Bill; No Dividends Levy. NEW YORK TIMES, A1 (September 21, 1950).


John D. Morris, Snyder Advocates a 75 Per Cent Tax on Excess Profits. NEW YORK TIMES, A1, Nov. 16th, 1950.

John D. Morris, Congress to Speed New Arms Billions; Profit Tax is ‘Sure’. NEW YORK TIMES, A1, Dec. 3rd, 1950.


10% Tax Surcharge Called Insufficient. NEW YORK TIMES, A19 (December 25, 1967).


Handler, M. S. 10% Income Surtax in Effect Today, Yielding $2.2-Billion Here. NEW YORK TIMES, A37, July 14th 1968.

Mills Sees Need to Extend Surtax. NEW YORK TIMES, A22, Aug. 15th 1968.

Wise, David. The Twilight of a President. NEW YORK TIMES, November 3, 1968.


Kraus, Albert L. The Surtax Debate. NEW YORK TIMES 59 (June 11, 1969).

Surtax Opposed by Most in Poll. NEW YORK TIMES, A14 (July 3, 1969).

Fraud on the Tax Front. NEW YORK TIMES, E14 (December 21, 1969).
“TO TAKE THE PROFIT OUT OF WAR”:
WORLD WAR II AND THE CREATION OF THE MODERN AMERICAN TAX REGIME

Joseph J. Thorndike
Director, Tax History Project at Tax Analysts
Scholar in Residence, University of Virginia

Draft of March 27, 2006

Preliminary Draft:
Please do not quote, cite, or circulate without author’s permission.
I. INTRODUCTION

Until World War II, only the rich paid income taxes. Lawmakers had traditionally confined the levy to the upper strata of American society, using it to balance broader, more regressive taxes on consumption. Critics decried the narrow focus. “When men once get the habit of helping themselves to the property of others,” complained the New York Times in 1909, “they are not easily cured of it.”¹ In fact, the editors were right: it took more than a quarter century for this rich man’s burden to reach the middle class. Only under the extraordinary fiscal pressure of World War II did lawmakers prove willing to transform the income levy from a class tax to a mass tax.

Between 1939 and 1945, Congress repeatedly lowered the threshold for taxable income, increasing the number of taxpayers sevenfold. (Figure 2.) By war’s end, more than 90 percent of American workers were filing tax returns, establishing the nation’s newest – and most pervasive – civic ritual. At the same time, lawmakers pushed rates skyward, with the top bracket eventually peaking at 94 percent. (Figure 3.) Together, these changes made the income tax a fiscal workhorse, boosting revenue from $1.0 billion in 1939 to $18.4 billion in 1945. By war's end, the tax was raising 40 percent of total revenue, displacing excise levies as the principal source of federal funds. (Figure 4.)²

The mass income tax was not the only wartime tax reform. The crisis-borne tax regime included a range of levies, including a corporate excess profits tax, an array of excise duties on consumer goods, and a small but growing flat-rate wage tax dedicated to social insurance programs. But it was the mass-based individual income tax that defined this remarkably durable tax regime. Its broad sweep, revenue buoyancy, and political resilience making it the centerpiece of federal finance for the rest of the twentieth century.

The transformation of the individual income tax arose from a confluence of factors, including politics, conviction, and necessity. Historians have emphasized the last, giving the story a functional and teleological cast. But while necessity was certainly the mother of this fiscal invention, ideas and ideology were vital, too. The modern income tax did not spring fully formed from the minds of wartime lawmakers. Nor was it ever a sure thing. Rather, it emerged from a vigorous debate over the meaning of tax justice and shared sacrifice.

II. BACKGROUND

The fiscal watershed of World War II followed a decade of bitter conflict over tax reform. The New Deal included a variety of important tax reforms, including steep new income taxes on the rich, heavy estate taxes, and a novel levy designed to regulate the distribution of corporate profits. But by the late 1938, conservative opponents of the New Deal had dug in their heels, thwarting plans for further innovation. By decade’s end, Roosevelt’s campaign for

---


progressive reform had become a rearguard action, with Administration officials fighting valiantly – if often vainly – to protect past victories. Nonetheless, New Deal tax debates left an enduring mark on the federal revenue system. Roosevelt’s commitment to progressive taxation helped transform political discourse, making progressivity a popular touchstone for fiscal reform. That political and rhetorical shift played a vital role in shaping the wartime tax regime.5

A. New Deal Conflicts Over Tax Fairness

Bitter struggles over taxation were a hallmark of the 1930s. Conservatives accused Franklin Roosevelt of trying to “soak the rich,” imposing steep taxes on income, estates, and corporate profits. Liberals complained that the system burdened the poor, taxing consumption too heavily and wealth too lightly. Perhaps most striking, advocates of “social taxation” – the creative use of tax policy to remake society and reform the economy – bemoaned the absence of innovation from Roosevelt’s fiscal agenda. At various points, FDR responded to these exogenous complaints, offering new initiatives and scuttling old ones. But ultimately, some of the most important arguments over tax policy occurred within the Roosevelt Administration, where two groups of presidential advisers squared off with very different notions of tax justice.6

Roosevelt’s tax advisors were a heterogeneous bunch, including lawyers and economists, planners and anti-trusters, Democrats and Republicans. Generally speaking, however, they fell into two groups. The first, populated by some of the Administration’s most senior lawyers, endorsed “high-end” progressivity: the notion that steep taxes on the rich should be used to

---

5 Brownlee, Federal Taxation in America, 112.
balance necessary, if unpalatable, taxes on the poor.

The second, which included most of the Administration’s professional tax experts, urged a program of “low-end” progressivity: the idea that new income taxes on the middle class could be used to finance a much-needed rollback in regressive consumption taxes.

Advocates of high-end progressivity acknowledged the regressive burden of federal consumption taxes. But they considered these taxes a necessary evil, especially as deficits continued to mount and Roosevelt struggled to defend himself against charges of fiscal recklessness. Advocates of high-end progressivity believed that heavy taxes at the top of the income scale would lend moral legitimacy to the revenue structure as a whole, including its more regressive elements. They argued for steep new levies on wealth and income. Even more important, they railed against tax avoidance by the rich, complaining that wealthy Americans were shirking their moral responsibilities to the nation and the fisc.

Fans of low-end progressivity cautioned against the punitive taxation of wealth and business, insisting that soak-the-rich taxes did a poor job of raising revenue. Instead, they endorsed a broad expansion of the income tax to the middle class, confident that such a levy could pay for a much-needed rollback in consumption taxes. These fiscal specialists understood the need for new social programs, and they were sympathetic to calls for sweeping economic reform. But they were wary of using taxes as a tool for reform. Better, they argued, to focus the revenue system on raising money, which could then be used to finance other, more ambitious items on the New Deal agenda.

---

7 The term “high-end” progressivity was not used by contemporaries in the 1930s, nor was its complement, “low-end” progressivity. I’ve used these terms, however, to more clearly delineate the policy preferences of these two, discrete groups, which while ideologically sympathetic with one another, were often divided over policy prescriptions.
The economists at Treasury spent much of the 1930s honing their case broad-based income taxation. But for most of the decade, the idea was something to warm the hearts of Republicans and a just a few conservative Democrats; liberals and moderates generally found it abhorrent. As a result, advocates of high-end progressivity carried the day, first with the president and later with his congressional allies. Between 1934 and 1937, Roosevelt pushed through a series of controversial tax reforms, including the Wealth Tax of 1935, the Undistributed Profits Tax of 1936, and the anti-loophole Revenue Act of 1937. These disparate reforms shared at least one principal objective: all were designed to shift the tax burden toward the rich, boosting the statutory and effective rates on wealth and income.

Roosevelt’s tax agenda outraged business leaders, Republicans, and even many conservative Democrats. By the mid-1930s, congressional opponents of high-end progressivity had marshaled their forces, and after 1937, they thwarted administration plans for additional tax innovation. They even managed to repeal the undistributed profits tax, FDR’s ambitious attempt to revamp the taxation of corporate earnings with an eye toward structural economic reform. As the Depression decade came to a close, the New Deal tax agenda – defined principally by its commitment to high-end progressivity – seemed little more than a distant, and disappointing, memory.

B. Planning for War

In fact, the legislative reverses of the late 1930s did not vitiate the political importance of New Deal taxation. While high-end progressivity lost ground on Capitol Hill, low-end progressivity continued to percolate through the federal bureaucracy, especially in the Treasury Department. Unable to make their case during the class-tinged debates of the 1930s, advocates of
low-end progressivity found an opening when lawmakers began struggling to pay for mobilization. Ultimately, Franklin Roosevelt and his advisers offer a hybrid version of progressive tax reform, uniting high- and low-end progressivity to champion a broad-based, steeply progressive personal income tax.

After 1937, a resurgent depression wreaked havoc with the federal budget, with tax revenues failing to keep pace with expenditures. Treasury Undersecretary Roswell Magill, a leading tax expert from recruited from Columbia University, asked his staff to devise a program that would raise at least $500 million. Carl Shoup, a liberal economist with impeccable academic credentials, responded with a plan for heavier estate and income taxes. Notably, however, he declined to endorse higher income tax rates on the rich – a mainstay of New Deal tax reform for most of the 1930s. Rates at the top were already too high, Shoup declared, discouraging work and investment; if anything, they should go down, not up. New revenue should come from the lower and lower-middle brackets of the income tax. Most of these taxpayers could afford to make a larger contribution, Shoup contended, and relative to their rich compatriots, they were substantially undertaxed.  

In 1939, another Treasury economist, Roy Blough, reached a similar conclusion. Steep rates on the rich were undesirable, he acknowledged, hampering investment and slowing the nation’s recovery from the depression. But high-end progressivity was also a political necessity. As he and Shoup concluded in a report they coauthored, “It seems most important to make the income tax a tax that will appeal to the community's sense of justice, even at the cost of having a

---

8 Carl Shoup, "Plans for Additional Revenue," in Box 43; Methods of Raising Additional Revenue; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1938).
higher rate scale.” 9 But the political capital derived from steep marginal rates should be used to advance plans for low-end progressivity. Income tax exemptions could be slashed dramatically without sacrificing progressive ideals, Blough maintained, and a broader tax base would provide ample new revenue to the cash-starved Treasury. The income tax was the fairest means of raising new revenue, Blough argued, allocating the tax burden according to an individual’s ability to pay. It was better than any sort of consumption tax, which would necessarily burden the poor more than the rich.

For most of the 1930s, broad-based income taxation found little support among congressional Democrats. Sen. Robert M. LaFollette, Jr., a maverick progressive Republican, championed the idea throughout the decade, and he had considerable support among his GOP colleagues. But liberals were cool to the idea, dimming its legislative prospects. “I personally doubt whether we can get this thing through called broadening the tax base,” Treasury Secretary Henry Morgenthau warned his staff in 1938. Better to focus on the rich and upper middle class, Morgenthau suggested, closing loopholes and ensuring that rich taxpayers were actually meeting their obligations. 10

Events conspired to rescue broad-based income taxes from political purgatory. As revenue shortfalls continued, Morgenthau warmed to the idea. So did Roosevelt, a longtime advocate of keeping the income tax narrowly focused on the rich. Morgenthau and Roosevelt began to listen when Treasury economists described their plans for a broader tax base, prompting

---


FDR to float the idea publicly in 1939. The income tax should be extended “a little bit,” he told reporters. “It wouldn’t bring in much revenue, but it does give added responsibility of citizenship.”¹¹ In fact, exemption cuts would have raised substantial revenue, which is why they were under consideration in the first place. But Roosevelt was careful to minimize the issue, if only to give Congress room to maneuver. Observers expected the exemption cut to unleash a firestorm of protest, and even Roosevelt questioned its viability.¹²

In fact, legislative prospects were better than Roosevelt knew. By the time Congress met to consider the first Revenue Act of 1940, soaring defense costs had transformed the political landscape. Congress began debating a variety of tax options, including a broad-based sales tax. But the individual income tax soon emerged as the most likely source of new money. Support for lower exemptions seemed to cut across the political spectrum, with even many liberals embracing the idea.

Revenue was almost always offered as the principal justification for lower exemptions. But as the war became a reality for Americans, the notion of sacrifice crept into the debate. In general terms, sacrifice had long been used to defend unpopular taxes. In 1932, for instance, congressional leaders had asked Americans to accept steep new taxes as a means to counter the

---


depression. “I do hope the patriotic American in this present emergency will be willing to undergo sacrifices for the common good of the people,” declared the chairman of the House Ways and Means Committee.\textsuperscript{13}

War made such rhetoric even more resonant, conflating fiscal sacrifice at home with mortal sacrifice on the battlefield – even before American troops entered the fight. “I have called for personal sacrifice,” Roosevelt told Congress in January 1941. “A part of the sacrifice means the payment of more money in taxes.” Roosevelt stressed the need for more revenue, but he refused to compromise his longstanding commitment to progressive reform; “the principle of tax payments in accordance with ability to pay,” he declared, “should be constantly before our eyes to guide our legislation.” For Roosevelt, sacrifice and progressivity were inextricably linked. If lawmakers were faithful to progressive ideals, then voters would respond nobly, “putting patriotism ahead of pocketbooks.”\textsuperscript{14}

Roosevelt’s high-minded call for patriotic sacrifice came with a negative counterpart. If self-abnegation was noble, then profiteering was reprehensible. “No person should try, or be allowed, to get rich out of this program,” the president insisted. Roosevelt believed that steep taxes were a vital weapon in the battle against profiteering, and many lawmakers shared his concern. Profiteering remained a fixture of fiscal debate throughout the war, influencing the taxation of both businesses and individuals.


III. WARTIME TAX REFORMS

In the early stages of the war, many tax debates focused on business taxation, with profiteering a source of worry and legislative determination. Soon, though, individual taxes joined the spotlight, their broad impact and political resonance a source of endless controversy. Revenue needs continued to drive most congressional tax debates, as policy makers struggled to finance a global war. But the politics of fairness and sacrifice remained vital to the structure of federal taxation, ensuring that income taxes would assume pride of place in the new, war-borne tax regime.

A. Taxing Business Profits

Business taxes were the subject of remarkable agreement in the opening stages of World War II. As U.S. officials began searching for ways fund new defense spending, most agreed that corporations would foot much of the bill. Vigorous arguments soon emerged over the details, particularly when lawmakers began crafting a new excess profits tax. But steep levies on corporate profits were broadly considered a vital, albeit temporary, element of national finance.

Early in his presidency, Franklin Roosevelt had endorsed heavy taxes on wartime profits. “The time has come,” he declared in 1934, “to take the profit out of war.”\footnote{Arthur Krock, "Move to End War Profit Tied to New Deal Timing," \textit{New York Times}, Dec 16, 1934.} In fact, the time had come, and gone, and come again. The excess profits tax of World War I, which had proven enormously effective as a revenue tool, was explicitly designed with moral considerations in mind. Political leaders had defended the tax not simply as a revenue device, but as a tool for ensuring fairness and social equity. The tax died a speedy death in the early 1920s, done in by its
political opponents, as well as its own complexity. But politicians of every stripe continued to pay lip-service to the steep taxation of wartime profits. Presidents Harding, Coolidge, and Hoover all endorsed the principle, as did the Democratic and Republican presidential platforms of 1928.16

Critics worried that excess profits taxes could be exploited by the enemies of private enterprise. If the profit could be taken out of war, wondered the Wall Street Journal in 1935, couldn’t it also be removed from depression, natural disaster, or any sort of national crisis?17 Indeed, many New Deal foes suspected Roosevelt of using the war profit debate to intensify his more general attack on corporate America and private sector prerogatives.

In fact, Roosevelt’s interest war profit taxation was sincere. He believed deeply that no one should profit from national misfortune, abhorring the juxtaposition of sacrifice abroad and self-indulgence at home. When Europe took the decisive plunge into full-scale war in late 1939, Roosevelt seized the initiative to revive his plans for steep excess profits taxation. Worried that his Treasury Department would resist a heavy profits tax, the president offered his own, very general plan. In July 1940, he asked Congress for a steep new levy on business profits, farming his request in resonant terms of shared sacrifice. “We are asking even our humblest citizens to contribute their mite,” Roosevelt declared. “It is our duty to see that the burden is equitably

16 Hurley’s tax would have imposed the tax on any corporate or individual income that exceeded a three-year average of pre-war income. "War-without-Profit," Time, 24 December 1934.

distributed according to ability to pay so that a few do not gain from the sacrifices of the many.”

Again, FDR was careful to marry sacrifice and progressivity in his call for wartime tax reform. Congress responded with the Second Revenue Act of 1940, an extraordinarily complex piece of legislation that included a new excess profits tax. Congressional leaders defended their handiwork. “Our taxes must follow the intricacies of business and not attempt to bend business to the pattern of simplicity we should all like to see in taxation,” observed Ways and Means Chairman Robert Doughton, D-N.C. Much of the bill’s complexity grew out of one issue: Which profits were, in fact, excess? What portion of a company’s income should be regarded as a product of the defense emergency rather than normal operations? Congress and the White House disagreed vehemently on this issue, which had enormous implications for the nature and operation of the excess profits tax.

Treasury experts proposed a tax that would define excess profits in relation to a “normal” percentage return on invested capital. This normal return might be arbitrarily fixed by government officials, or it might be derived from the prewar returns enjoyed by a taxpaying company. Many, but not all, tax experts preferred this method. By contrast, congressional

---


leaders favored the “average-earnings” method of measuring excess profit. Under this scheme, earnings during the taxable, wartime year that exceeded the average earnings during a three-year, prewar base period would be subject to the new tax.\textsuperscript{21}

Eventually, the House and Senate agreed to let taxpayers choose their preferred method, subject to certain rate penalties for those selecting the pre-war profits technique. It was a business-friendly decision. Companies with big profits in the averaging period could choose the pre-war earnings method and keep making big profits during the war. Corporations making small profits in the averaging years could choose the invested capital method and preserve their shot at a good return. Meanwhile, a large exemption left most companies entirely exempt. Clearly, sacrifice had its limits: lawmakers were eager to allow room for business to prosper.\textsuperscript{22}

Just days after Congress completed work on the second Revenue Act of 1940, Treasury Secretary Morgenthau gave lawmakers a good scare when he made the casual observation that any business should be satisfied with a 6 percent return on invested capital. To wary ears, this sounded like a trial balloon for a 100 percent rate above that amount. Treasury officials soon backed off this statement, and Morgenthau later claimed it was a slip of the tongue. But in fact, rates would approach that level within just a couple of years. In 1941, lawmakers raised the

\textsuperscript{21} See, for example, John Richard Hicks, Ursula Kathleen Webb Hicks, and L. Rostás, \textit{The Taxation of War Wealth}, 2nd ed. ed. (Oxford: Clarendon Press, 1942).

\textsuperscript{22} The Second Revenue Act of 1940 raised the normal income tax rate for corporations to 24 percent for all companies earning more than $25,000 in net income. Excess profits were taxed at rates ranging from 25 percent to 50 percent, depending on income. In computing the new tax, companies were granted a flat $5,000 exemption, plus a credit computed by one of the two methods described above. The bill also retained the accelerated amortization first proposed by the White House, although it notably deprived the government of any role in the postwar disposition of these facilities. See Roy G. Blakey and Gladys Blakey, "The Two Federal Revenue Acts of 1940," \textit{American Economic Review} 30, no. 4 (1940): 733.
excess profits rates across the board, establishing a range from 35 percent to 60 percent. They also tightened available deductions to further boost revenues.\(^{23}\)

In 1942, lawmakers abandoned the graduated rate structure for the excess profits tax, imposing a flat 90 percent rate. In 1943, the rate climbed again to 95 percent, although the exemption grew dramatically as well, rising from $5,000 to $10,000. Such hikes were further offset by a range of relief provisions, enacted at several points during the war. Designed to ease inequities and encourage certain types of business activities, these provisions brought forth a deluge of refund claims from almost 10,000 corporations. Ultimately, such claims cost the government about $388 million. Various other provisions also softened the blow to corporate treasuries. Caps on the total amount of income that might be collected from a corporation (80 percent of total profit) and a 10 percent postwar refund on taxes paid (which gave the tax an element of compulsory lending) reduced the effective rate to as little as 72 percent.\(^{24}\)

Throughout the war, Treasury officials continued to argue for the invested capital method of calculating excess profits. In May 1941, the assistant secretary for tax policy complained that the availability of both methodologies was eviscerating the tax; one company with $70 million in war orders had paid no excess profits tax whatsoever. Clearly, he declared, the tax was failing in its stated purpose: some companies were making enormous profits from the war. Roosevelt himself joined the argument, calling for more stringent rules. Companies must be forced to share


\(^{24}\) Treasury supported some of these relief provisions, notably the 80 percent ceiling. See Henry Morgenthau, "Statement of Secretary Morgenthau before the Ways and Means Committee of the House of Representatives," in *Box 34; Defense and War; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD.* (1942); Paul, *Taxation in the United States,* 320-21.
the burden of wartime sacrifice. “Excessive profits undermine unity and should be recaptured,” he declared in his January 1942 budget message. “The fact that a corporation had large profits before the defense program started is no reason to exempt them now.”

But Congress stood firm, insisting on the two-method system. This congressional intransigence may have reflected postwar intentions; an invested capital standard could have been converted to peacetime use, but the average earnings technique was clearly a wartime device. Similar arguments had surrounded the excess profits tax of World War I. But this time, most observers expected the excess profits tax to be a short-term addition to the tax system. Even those who favored peacetime use of the levy were skeptical of its long-term viability. In 1942, Congress made explicit its intention to abandon the tax once peace returned.

Excess profits taxes were not the only business levies imposed during the war, although they were certainly the most important. Congress repeatedly raised regular corporation income taxes. The First Revenue Act of 1940 raised rates, introducing three brackets of 15 percent, 17 percent, and 19 percent. The top rate applied to corporations with more than $25,000 in net income, and a special 10 defense surtax raised the effective rate even higher. In 1941, lawmakers raised the normal corporate rate from 24 percent to 31 percent and introduced a new surtax of 6 percent on net income under $25,000 and 7 percent on amounts over that threshold. Effectively, rates topped out at 24 percent for large companies, while those with smaller incomes paid 15 percent to 19 percent. In 1942, rates went higher still. The revenue from this tax was substantial,

---


ranging from $3.7 billion in fiscal 1941 to $4.5 billion in 1943. But the tax raised far less than
the levy on excess profits, which ranged from a $3.4 billion in 1941 to $11.4 billion in 1943.27

Despite its various shortcomings, the excess profits tax enjoyed grudging acceptance
throughout the war. As early as 1940, political observers noted its popularity. Calls for a profit
tax “were representative of a widespread determination that the misfortunes of war should not be
taken advantage of to create a lot of millionaires,” reported economist Roy Blakey, one of the
leading observes of federal tax policy.28 The levy was a political necessity, bolstering fairness at
a critical point in the nation’s political and fiscal history. “It was an answer to those who sought
insurance that none would profit from the nation’s misfortune,” wrote liberal economist Harold
Groves. “It was demanded as a monetary counterpart to the sacrifice being made by persons who
entered the armed forces.” Its essential fairness, as well as its productivity, silenced many critics.
“The profits tax, as a war measure,” Groves concluded, “stems from some sound roots in the
soils of expediency and principle.”29

Even business leaders accepted the excess profits tax, at least in theory. “To the best of
my knowledge, no business executive opposes the principle of an excess profits tax which seeks
to return to government exorbitant profits arising directly or indirectly from the war effort,”
wrote tax lawyer George Douglas. Leaders of the private sector had numerous complaints about
the operation of the tax, and Congress worked diligently throughout the war to ameliorate some
of the more egregious problems. But the excess profits tax enjoyed remarkable, if not always

enthusiastic, support from most business leaders. Even the National Association of Manufacturers endorsed a 90 percent levy in 1942.30

Still, business was restive under the burden. Randolph Paul complained that an initial burst of patriotism, business enthusiasm faded dramatically over the course of the war. And certainly, business leaders were vocal in their support for moving more of the tax burden onto individuals rather than businesses. Often, they used inflation to bolster their case, cloaking their quest for lower taxes in the mantle of macroeconomic responsibility. But it was abundantly clear to most observers that many business groups were principally interested in limiting their own tax burden and shifting it to consumers.

B. Taxing Individuals in Total War

The broad-based, individual income tax was the defining characteristic of the World War II tax regime. Fashioned jointly by Congress and the Roosevelt Administration, it was a testament to compromise and cooperation. But the transformation of the income tax – its conversion from a “class tax” to a “mass tax” – was never a foregone conclusion. Many politicians and opinion leaders preferred a national sales tax, insisting that it would raise much-needed revenue while also curbing inflation. Even some members of the Roosevelt Administration agreed.

But Franklin Roosevelt rejected the sales tax. The levy, he declared, was inefficient, unnecessary, and unfair. In its search for revenue, Congress should look first to wealth and big

business, he maintained. And when it came time to tax Americans of modest means, the income tax was better suited to the task. Roosevelt was willing to tolerate a range of narrow consumption taxes, especially on luxuries. But he refused to consider a general sales tax on personal consumption. That sort of levy, he complained, would spare the rich and soak the poor.

Roosevelt’s vigorous leadership derailed the drive for sales taxation. Lawmakers turned, instead, to the individual income tax, pushing it deep into the middle class. By war’s end, more than 90 percent of American workers filed annual returns. Fully 60 percent found they owed money to Uncle Sam, giving this rich man’s tax a newly democratic cast. At the same time, lawmakers acceded to Roosevelt’s demand that the personal income tax remain steeply progressive. Marginal rates soared above 90 percent for the nation’s wealthiest taxpayers, while effective rates peaked at almost 60 percent. These changes – representing a marriage of high-end and low-end progressivity – made the income tax a fiscal workhorse, the principal source of federal revenue. It would retain that distinction for the rest of the twentieth century.

1. Inflation

The transformation of the individual income tax was driven by revenue imperatives. The exemption cuts enacted between 1940 and 1943 were designed to raise money: lots of it, and very quickly. But unlike business levies, the income tax was also designed to fight inflation. By

---


32 To some degree, inflation also shaped the debate over corporate taxation, but the effect of business taxes on the overall price structure was the subject of much head-scratching among tax experts of the 1930s and 1940s. Most economists believed that corporate taxes had a modest influence on inflation, both for good and ill. See, for example, Groves, "An Appraisal of the Excess Profits Tax from a Fiscal Standpoint," 135-36; J. Wilner Sundelson, "The Federal Revenue System: Taxation During Inflation," in Box 62: Tax Reform Programs and Studies; Records
early 1941, Henry Morgenthau was deeply concerned by the prospect of rising prices. At his behest, Treasury tax experts had begun debating the best means to check the price spiral. They agreed on the importance of slowing consumer demand, embracing a range of taxes designed to drain consumers of their purchasing power.33 As the department’s tax experts later recalled in a fit of self-congratulation: “The record shows that the Treasury in the autumn of 1941 was practically alone, among responsible groups in the executive and legislative branches, in urging special anti-inflationary taxes.”34

In part, the Administration’s slow start can be explained by Roosevelt’s reluctance to consider new taxes on the poor and middle class. He did not share the Treasury’s inflation concern through the early years of the defense program. Convinced that the economy had room to absorb new government spending, he assured Congress that the sky was not, in fact, falling in. “I am opposed,” Roosevelt declared in January 1941, “to a tax policy which restricts general consumption as long as unused capacity is available and as long as idle labor can be employed.”35 But soon enough, events forced Roosevelt to reconsider his sanguine attitude.

Prices rose steadily throughout 1941, and by early 1942, a sense of crisis gripped policymakers.


34 “Treasury Anti-Inflation Tax Proposals,” in Box 34; Inflation, Depression, Recovery; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1942).

“Inflation was no longer something that might hit sometime in the nebulous future,” recalled Treasury General Counsel Randolph Paul. “It had gained a tight hold on the national economy.” In September, Leon Henderson, chief of the Office of Price Administration, reported a 17 percent price hike over the previous year. In December, he predicted another 20 percent rise within four months. Retail process were already increasing at 1.5 percent a month.\footnote{“Henderson Sees ‘Mild Inflation,’” \textit{New York Times}, 18 September 1941, 11; “Henderson Sees Living Cost Up 20\% By Spring and Urges Price Control,” \textit{New York Times}, 3 December 1941, 16.}

As Treasury inflation hawks were quick to point out, massive defense spending almost guaranteed a price spike. Wages would grow as production increased and labor supplies dwindled. Demand for consumer goods would rise in tandem, as newly flush workers found money burning a hole in their pockets. Meanwhile, defense production would crowd out civilian goods, leaving shelves bare and consumers agitated. The result? A vicious spiral of rising prices, as growing incomes chased a shrinking supply of available goods. Inventories might hold things in check for a while, but soon enough, shortages would be common. The outlook was sobering.\footnote{Paul, \textit{Taxation in the United States}, 284.}

As the federal government assumed a true war footing, the anti-inflation campaign developed rapidly and dramatically. Its most famous components, including price controls and rationing, were implemented in 1941 and early 1942. But tax experts believed that such controls were inadequate. Only by draining consumers of their new purchasing power could the nation avoid a ruinous price rise. Tax revenues had been going up for several years, thanks to a growing national income. Steeper surtaxes imposed in 1940 and 1941 had further increased receipts. But existing taxes were inadequate to the Herculean task of checking inflation. “Increased income lay
smoldering in millions of pockets,” wrote Randolph Paul. “It was an explosive factor in the economy.”

2. Taxing Consumption

Consumption taxes were a proven revenue tool, and many observers believed that they were also the best weapon against inflation. And since many consumption taxes were already on the books, they enjoyed a certain degree of popular acceptance. For most of the nation’s history, taxes on consumption had been the principal source of federal revenue, whether levied through the tariff or various excise duties. Even after the introduction of the income tax, consumption taxes remained a key source of revenue, propping up the federal tax system when income tax receipts fell precipitously during the Great Depression. It was only natural, then, that policymakers would look to consumption taxes as a key component of the wartime revenue regime. But which kind? While some policy makers, including those in the Treasury, favored excise taxes, others called for a national sales tax.

a. Excise taxes

In the spring of 1941, Assistant Treasury Secretary John L. Sullivan asked Congress for more than $1 billion in new consumption taxes, chiefly from various excise duties on consumer goods. Sullivan acknowledged that most excises “fail to meet the test of equity.” Indeed, Treasury economists had been arguing for more than a decade that excise revenue should be replaced by money from a broader income tax. This was the central pillar of low-end

---

38 Ibid., 283.
progressivity. But Secretary Morgenthau was determined to finance two-thirds of the war with taxes, not borrowing, and that ambitious goal demanded a flexible, powerful, and heterogeneous tax system. Consumption taxes were reliable and productive, making them an indispensable part of the wartime regime.39

Sullivan pointed out that Treasury had eschewed new taxes on necessities. Instead, the department’s experts had recommended levies on luxury items, including several that were broadly popular among all income groups. Taxes on candy, chewing gum, and soft drinks might not meet the traditional definition of luxury, but they were certainly non-necessities. As such, they could be taxed more heavily without unduly burdening low income Americans. Consumer sacrifice would be minimal.

The Joint Committee on Internal Revenue Taxation (JCIRT) responded to the Treasury plan with an even longer list of proposed consumption taxes. Chief of Staff Colin Stam told lawmakers that ability-to-pay taxes were already bearing too much of the burden, and a quick look at receipts revealed plenty of room for higher consumption duties. He added 14 excise hikes to the Treasury’s list, including levies on coffee, cocoa, sugar, and other common foodstuffs. Such levies involved more sacrifice, but certainly nothing intolerable.40

The Revenue Act of 1941 ultimately included some $846 million in new excise taxes. Some were specifically aimed at goods that competed with war production, including refrigerators and automobiles. Others fell on luxury items like jewelry, furs, and toiletries. A few

39 Committee on Ways and Means, Revenue Revision of 1941, 77th Congress, 1st Session, 24, 28-30 April, 1-2,5-9 May 1941, 47.
40 Ibid., 82-88. Stam also suggested several shifts among “ability” taxes, placing more reliance on corporate excess profits taxes and less on the individual income levies.
quasi-luxuries also made the list, such as playing cards, safe deposit boxes, radios, and club dues. (Some items in this latter collection, such as safe deposit boxes, were also considered vaguely progressive, since they were thought to burden affluent taxpayers). Alcohol taxes, as well as those on tobacco, always seemed to get free pass in fairness debates, their status as “sin taxes” insulating them from complaints about their regressive incidence. In 1941, lawmakers raised existing taxes on wine and distilled spirits, but not those on tobacco. Finally, even churches were asked to bear some of the new tax burden, albeit indirectly; entertainment programs offered by schools, churches, and scientific societies, once free from the admissions tax, were now made taxable.\footnote{U.S. Public Law 77-250, 77th Congress, 1st Session, 20 September 1941. See also Blakey and Blakey, "The Revenue Act of 1941," 821.}

In 1942, The Treasury recommended $1.3 billion in further excise hikes, including new or higher taxes for liquor, gasoline, cigarettes, soft drinks, candy, and chewing gum. As always, the Treasury acknowledged that such taxes were regressive. But they had a few redeeming qualities, Morgenthau told lawmakers. Several, such as the tax on gas, were levied on scarce items; these were designed as much to forestall shortages as they were to raise revenue. Others fell on items that were widely consumed but clearly unnecessary: candy and chewing gum, for instance. Because demand was fairly inelastic for most of the targeted items, the taxes would cause minimal disruption. “Needed revenue will thus be obtained, consumer purchasing power will be tapped, the producers will not be injured, and the consumers will not be taxed on the necessities of life,” Morgenthau said.\footnote{Committee on Ways and Means, Revenue Revision of 1942, 77th Congress, 2nd Session, 3, 5, 9-13, 16-20, 23-24 March 1942, 7.} After the usual wrangling, Congress approved the
Revenue Act of 1942 with more than $650 million in new excise revenue, including a major hike in liquor and tobacco taxes.\textsuperscript{43}

In 1943, another round of excise increases continued the upward march. Morgenthau recommended $2.5 billion in new consumption taxes; Congress agreed to just $1 billion, as lawmakers began to resist incessant Administration demands for additional revenue. In 1943, Congress raised another billion from the same source. Over the course of war, excise taxes boosted the government’s bottom line quite handsomely. Consumption revenue climbed from $2.5 billion in fiscal 1941 to $6.2 billion in 1945.

At the same time, the excise yield declined relative to other revenue streams. Producing more than 29 percent of federal revenue in fiscal 1941, excises provided just 13 percent in 1945. Newer, more productive taxes had eclipsed these venerable revenue tools, and they would never again regain their prominent place in the federal tax system. Indeed, many would disappear entirely in the decade after the war.\textsuperscript{44}

b. Sales Taxes

Excise levies were the most obvious tool for taxing consumption, but they were not the only one. During every wartime debate over excise taxes, a controversial alternative loomed in the background: a general sales tax. Many policymakers preferred a sales tax, imposed either on manufacturers or retailers. At various points in the early years of the defense program, it seemed that such a tax was a near certainty. In the spring of 1942, unhappy Treasury officials – who

\textsuperscript{43} U.S. Public Law 77-753, 77\textsuperscript{th} Congress, 2\textsuperscript{nd} Session, 21 October 1942. See also Roy G. Blakey and Gladys Blakey, "The Federal Revenue Act of 1942," \textit{American Political Science Review} 36, no. 6 (1942): 1075-76.

\textsuperscript{44} United States Bureau of the Census, \textit{Historical Statistics of the United States, Colonial Times to 1970}.
urgently resisted sales taxation at every turn – were resigned to the idea. Ultimately, however, President Roosevelt deflated the drive for sales taxation by taking a firm and very public stand against the idea. With strong support from organized labor and compelling case from his Treasury experts, FDR ensured that income taxes would be the principal levy for taxing average Americans.

Sales tax advocates insisted that a broad–based consumption would raise ample revenue while also discouraging consumption; since these were the twin imperatives of wartime finance, what tax could be better suited to occasion? The Washington Post and the Wall Street Journal regularly endorsed the idea, and even the New York Times expressed cautious interest. But the Treasury was unconvinced. While the department’s Division of Tax Research prepared several plans for a national sales tax, all were prompted by congressional requests. The division’s leadership and staff were firmly opposed to the idea. In their view, numerous alternatives were preferable, including targeted excises and a much broader income tax. Most congressional Democrats were inclined to agree.

Over time, the pressing need for revenue seemed to weaken the resolve of some sales tax opponents. Sen. Josiah Bailey, D-N.C., opening the door just a crack, told reporters that he would endorse a general sales tax “only as a last resort.” Senate Majority leader Alben Barkley, D-Ken., mused that a sales tax might be unavoidable. Even Ways and Means Chairman Robert Doughton, D-N.C., predicted in 1940 that a sales tax was in the offing. Doughton, who had made
a name for himself by leading a rebellion against the sales tax in 1932, told colleagues to expect one in 1941.  

Meanwhile, most business leaders kept up a steady drumbeat for the sales tax. In testimony before Congress, they repeatedly endorsed the tax as a vital revenue tool and efficient weapon against inflation. They also stressed the inequities of targeted excises. Narrow levies burdened unlucky items while leaving others unscathed. Often as not, decisions about what to tax were driven by political considerations, not any broader sense of fairness; well connected industries could avoid new taxes while less favored ones could not. The National Association of Manufacturers encouraged lawmakers to explore the possibility of a general sales tax; their timid endorsement of the idea in 1941 stressed the need for revenue and the inadequacy of existing levies. The New York Chamber of Commerce offered a similar proposal with a similar lack of enthusiasm; the explosive potential inherent in such ideas seemed to breed caution in these and other business groups.  

But popular attitudes toward sales taxation were hard to gauge. Opinion polls seemed to indicate substantial support. A survey sponsored by NAM in January 1941 found 32 percent of respondents supporting the idea. A year later, the Gallup polling organization found support among 47 percent of respondents. And in 1943, Gallup reported that fully 54 percent of Americans preferred a new sales tax to higher income levies.  

---

46 Revenue Revision of 1941, 55.  
Congress initially resisted the drive for a sales tax, but by 1942, the campaign had reached a fever pitch. As Congress prepared to tackle the most sweeping revenue bill in the nation’s history, business groups bombarded Capitol Hill with calls for a retail sales tax. Gone was the hesitancy of a year earlier; advocates were eager to declare their support. Most urged that a sales tax be used to forestall further hikes in personal income taxes. Others justified the innovation as a replacement for the motley collection of excise taxes. Petitions and letters piled up in congressional offices, and pro-sales tax spokesmen jammed the witness tables at committee hearings. Advocates insisted that a modest sales tax would raise enormous revenue. NAM predicted than an 8 percent tax would produce at least $4.4 billion annually. The National Retail Drygoods Association called for retail sales levy, pointing out that such a position was pretty noble for a bunch of storeowners.48

Congress seemed to be listening. Throughout the spring of 1942, press reports indicated growing support for some sort of sales tax. House and Senate leaders seemed open to the idea, while the two taxwriting committees were reported to be well along in their planning for such a tax.49 Talk of a sales tax grew so agitated that some observers managed to convince themselves


that White House was behind it all. “A Federal sales tax is in the works,” predicted *Time* magazine. “The entire body of New Deal thinking, which long opposed sales taxes as a burden on the poor, has switched completely.”

In fact, the Roosevelt Administration was still adamantly opposed to the sales tax. From the Oval Office to the Treasury, no Administration official wanted anything to do with it. Together, political leaders and technical officials marshaled a powerful campaign to discredit the idea. Treasury experts had long insisted that general sales taxes were unfair, unnecessary, and unworkable. When levied on consumer goods, a sales tax burdened necessities and luxuries alike. Excise taxes were easier to target at appropriate goods, ameliorating the regressive quality inherent in more general consumption taxes. On April 14, 1942, Roosevelt asked the Treasury to prepare some talking points on the sales tax, with an eye toward deflecting the idea.

Economist Roy Blough obliged him with a memo on the “Evils of the Sales Tax” that stressed its inequities, while scoring the levy for its heavy administrative burden.

While sales taxes promised to raise more revenue than selective excises, they were much less productive than many advocates suggested. “The sales tax advocates were romantics about

---


51 See, for example: Carl Shoup, ”The Federal Revenue System: The Sales Tax,” in *Box 62; Tax Reform Programs and Studies; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD.* (1934).

52 Lazlo Ecker-Racz, ”Some Considerations Respecting the Advantages of Increases and Additional Selective Excise over a General Sales Tax,” in *Box 1; Excise and Sales Taxes in General; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD.* (1941).

53 Roy Blough, ”Evils of the Sales Tax,” in *Box 44; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO.* (1942).
the amount of revenue the tax would yield,” remembered Randolph Paul. While ostensibly levied on all consumer expenditures, the tax actually drew from a much smaller base. Out of more than $80 billion in total annual spending during the war, Paul estimated, roughly $30 billion went to services, which were considered un-taxable. An exemption for food reduced the base even further, to $30 billion. To raise $5 billion in revenue, then, lawmakers would have to approve a tax of 17 percent—far more than even the levy’s champions would support.54

Organized labor opposed the sales tax, repeatedly and emphatically. In 1942, C.I.O president Philip Murray denounced the idea as “shocking,” calling instead for a modest expansion of the income tax and steep additional levies on wealth and business income. Murray stated his feelings in bold terms. “In peacetime a sales tax is vicious enough,” he declared. “but in wartime, when we are trying to assure our war workers of sufficient funds to maintain themselves, the proposed sales tax levy would be the equivalent of a military defeat.”55

But still, the pressure for sales taxation continued to mount. JCIRT experts offered several plans, and under duress, the Treasury delivered one as well. At least two Democrats on the Ways and Means Committee were sold on the idea. On the Senate Finance Committee, Sens. Arthur Vandenberg, R-Mich, and Harry Byrd, D-Va., mounted a bipartisan campaign for the tax. The committee chairman reported that more than half his members were sympathetic.56

54 Paul, Taxation in the United States, 325. In 1942, the Treasury predicted that a 10 percent retail sales tax exempting food would raise $1.7 billion; "Federal Manufacturers', Wholesale, and Retail Sales Taxes,” in Box 44; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO. (1942). See also Ecker-Racz, "Some Considerations Respecting the Advantages of Increases and Additional Selective Excise over a General Sales Tax."


56 “Sentiment Grows for the Sales Tax.”; Young, "General Sales Tax Considered Reluctantly by House Group."; Robert De Vore, "Treadway Predicts Sales Tax Adoption," Washington Post, Apr 7, 1942; Cotten, "House
Franklin Roosevelt never wavered. At various points in 1940 and 1941, he had seemed to leave the door open, if only barely. War exigencies, he suggested, made it prudent to leave all options on the table. But his refusal to disavow the idea was misleading; when the time came to declare his position, Roosevelt came out firmly against the sales tax in all its forms. Time and again, when asked by reporters to reconsider his position, the president stood firm.\(^{57}\)

Administration officials were not optimistic about their ability to stop the sales tax drive. In the spring of 1942, internal memoranda from the Treasury’s tax division took some sort of sales tax as a given.\(^{58}\) But in fact, presidential opposition, combined with pressure from labor groups, had tipped the balance of political opinion. It was a near thing, especially in the Senate Finance Committee, where the idea had real momentum. Randolph Paul later remarked that the absence of a sales tax from the Revenue Act of 1942 was one of the most striking aspects of the law. When forced to make a choice, however, lawmakers lined up with the president, choosing to focus on income, not consumption taxes. It was a fateful decision, shaping the federal revenue system – and the American economy – for decades to come.\(^{59}\)


\(^{59}\) "A Scheme for a Progressive Sales Tax," in *Box 44: Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO.* (1942); "Majority Backs Levy on Sales, George Says."
c. The Spendings Tax

While Roosevelt ultimately won his battle over sales taxation, the fight was long and hard. Congress was in a contrary mood, and Administration ideas were often rejected out of hand. *Time* magazine captured the atmosphere when it described a Treasury appearance before the Finance Committee in 1942. “Henry Morgenthau and his tax experts marched up Capitol Hill and marched right down again,” the magazine reported. “They came up to propose a new tax program; for all political purposes they were almost kicked downhill.”

*Time* was describing the ignoble fate of the most innovative tax proposal to come out of the wartime Treasury: a progressive spendings tax. Conceived to bridge the gap between advocates of sales and income taxation, the idea went precisely nowhere. Lawmakers dismissed it immediately, and the tax made scarcely a ripple in its own time, slipping beneath the turbulent waters of a busy legislative season.

But the spendings tax was nonetheless important, demonstrating the Roosevelt’s Administration’s commitment to tax fairness, even when it came to taxing consumption. Morgenthau offered his plan for the spendings tax on September 3, 1942, while Congress was well into its work on the huge revenue bill ultimately passed in October. His motives were two-fold. First, he faced an urgent need for money. Expenditures for fiscal 1942 were expected to exceed $80 billion, while revenue projections for the 1942 act were running at only $24 billion. Second, Treasury officials were still seeking new ways to control inflation. "The Treasury,"

---

Morgenthau asserted, "is seeking in these proposals to attack the problem at its roots and to attack it drastically."\(^{61}\)

The spendings tax was designed to supplement, not replace, the individual income tax. In fact, Treasury suggested that the two taxes could be administered jointly, with taxpayers filing a combined return and sending in a single payment. Individuals would calculate their total spending indirectly, subtracting savings from the total amount of available funds (including current income and reductions in capital). Savings were broadly defined to include debt repayment, life insurance premiums, purchases of capital assets, gifts and contributions, tax payments, and increases in bank balances.

Using this base, the spendings tax would then be imposed in two parts: a flat rate tax of 10 percent to be refunded after the war, and a progressive surtax. The refundable portion, which would have applied to all individuals already paying the income tax, assessed a flat tax on all spending; the proposal made no provision for any deductions, nor did it exempt any minimum level of spending. Because it was refundable, this portion of the tax amounted to a compulsory, non-interest loan from the taxpayer to the government. The surtax, by contrast, was to be imposed at progressive rates ranging from 10 percent to 75 percent. All expenditures in excess of certain exemption levels would be subject to the tax. Treasury left open the possibility that various "extraordinary expenditures" might be made deductible.\(^{62}\)

Treasury officials developed the spendings tax with a close eye on its alternatives, principally a further increase in income tax rates or the enactment of a general sales tax. The


\(^{62}\) Ibid.
spendings tax enjoyed several advantages over the former, according to Treasury studies. First, it would more effectively curtail consumer spending, making it a better weapon for fighting inflation. Second, its progressive rate structure, if sufficiently steep, might permit fairly close regulation of individual spending levels. Third, the spendings tax was thought to be more politically palatable than further extension of the income tax; because individuals had considerable discretion in determining their spending patterns, they could largely determine their own tax liability, making the spendings tax less onerous than income levies.63

Treasury officials also believed the spendings tax enjoyed several important advantages over a general sales tax. First, it allowed for a much more efficient means of granting personal exemptions. Second, it would be less inflationary than a sales tax. Third, it would not upset wartime price controls, then in their early stages and already under considerable strain. Sales taxes, by contrast, were thought to involve such problems, given their tendency to raise the cost of production. Finally, sales taxes were considered more administratively burdensome, demanding an entirely new collection structure, while the spendings tax would piggyback on the income tax.64

But what appealed most strongly to advocates of the spendings tax was its element of choice. The tax, noted Randolph Paul, would have “left the taxpayer substantial latitude and freedom.” People who wanted to spend lavishly were free to do so – they would simply have to pay the price in extra taxes. If, on the other hand, an individual wanted to minimize his tax

63 “Proposal for a ‘Consumption Expenditure Tax’,” in Box 6; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO. (1942).

64 “The Spendings Tax,” in Box 6; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO (1942).
burden, then he could adjust his consumption accordingly. “Thus to a considerable extent,” Paul concluded, “he would have been his own tax assessor.”

Congress was unmoved, complaining that the tax was too complex. The Senate Finance Committee dispatched it immediately, and few outside the Treasury lamented its passing. Robert C. Albright of The Washington Post offered the best epitaph, calling it "Morgenthau's morning glory.” As he observed: “It opened Tuesday morning and it folded before noon.”

3. Taxing Income
The transformation of the individual income tax – including the expansion of its base and the increase in its rates – unfolded over several years. Lawmakers slashed exemptions repeatedly, bringing millions of new taxpayers into the system. (Figure 1) At the same time, they raised rates across the board. For many political leaders, including President Roosevelt, this was the vital moral tradeoff behind wartime taxation. New taxes on the poor and middle class had to be balanced by steep new levies on the rich. This bargain worked itself out in debates over the relative importance of consumption and income taxes. But it also shaped the structure of the income tax itself, producing a tax both broad and steep.

65 Paul, Taxation in the United States, 294.
66 Robert C. Albright, “Gallery Glimpses,” Washington Post, 6 September 1942. Over the next year or so, Treasury did continue to study the levy, convinced of its virtues. But it never again received serious consideration.
Once again, revenue needs were the driving force. But the levy was not simply a product of exigency and expediency. It was also the realization of thoughtful design and careful planning. The mass income tax, while deficient in any number of ways, represented a triumph for low-end progressivity and its longtime advocates in the Treasury Department. At the same time, it featured elements of high-end progressivity, with steep marginal rates on the rich used to justify new taxes on the middle class.

In the first six months of 1940, German armies swept across Northern Europe, prompting Roosevelt to send Congress a series of military spending requests, each bigger than the last. On May 16, the president asked for $1.2 billion; on May 30, he asked for another $1.3 billion. As lawmakers cast about for ways to pay those bills, the Treasury urged a cut in income tax exemptions, again invoking the need for sacrifice. “I am convinced that the public is willing and
ready to accept the personal sacrifices of paying the additional taxes that are necessary to provide the country with adequate national defense,” Morgenthau told Congress. In fact, unsolicited contributions had been arriving at Treasury for weeks, with Americans volunteering amounts from 10 cents to $500.\(^6^8\)

The First Revenue Act of 1940 cut exemptions by a fifth; individuals were taxed on income over $800, rather than $1,000, while couples paid on anything over $2,000, rather than $2,500. Lawmakers expected the cut to create two million new taxpayers and to raise about $75 million. Treasury acknowledged that the lower threshold for returns would raise enforcement costs; each return cost between 50 cents and $1.56 to process, depending on its complexity. But officials believed that lower exemptions would also boost compliance among those already liable under the higher exemptions. Jolted into awareness by the new exemption cuts, taxpayers previously on the margin of taxability would probably pay up, offsetting the $8 million in new administrative costs.\(^6^9\)

The 1940 law also raised surtax rates across the board. The biggest hikes were in the lower and lower-middle brackets, just as Shoup and Blough had suggested in the late 1930s. Taxpayers earning between $6,000 and $100,000 also saw major increases. The rate hikes were expected to raise $177 million. Finally, the law imposed a special, 10 percent “defense surtax”

---

\(^6^8\) Committee on Ways and Means, The Revenue Bill of 1940, 76th Congress, 3rd Session, 10 June 1940, 7; Committee on Finance, Revenue Act of 1940, 77th Congress, 1st Session, 12-14 June 1940, 33.

\(^6^9\) Revenue Act of 1940, 32. Only $14 million of the new revenue would come from the new taxpayers; the other $61 million would result from existing income taxpayers, who would see their liability rise as the exemption fell. In addition, Treasury believed that better compliance among existing taxpayers would raise perhaps $25 million.
on almost every existing internal revenue levy. This “supertax” was slated to expire in five years, while the exemption cuts and regular rate hikes were permanent.\(^{70}\)

The income tax revisions of the 1940 bill were surprisingly uncontroversial. Almost no one objected to the exemption cuts or the rate hikes. To be sure, business groups warned that steep surtaxes diminished the incentive to work, and labor groups complained about any broadening of the tax base. But most people swallowed hard and accepted the heavier taxes with minimal complaint.

In 1941, things got more contentious. House lawmakers recommended another round of rate hikes. They did not, however, propose a further cut in exemptions. In April, Treasury officials had testified that exemptions were already low enough; individuals earning as little as $17 a week were paying the income tax. Republicans had needled the Roosevelt Administration on this point, asking if civilians were sacrificing as much as military personnel. But Assistant Secretary Sullivan defended existing exemptions as a matter of fairness and practicality.\(^{71}\)

The president, however, had other ideas. On July 31, he wrote Ways and Means Chairman Bob Doughton to endorse new, lower exemptions: $750 for individuals and $1,500 for couples. While revenue needs certainly prompted FDR’s request, he again offered a moral, not an economic argument for the broader tax. “I am convinced that the overwhelming majority of our citizens want to contribute something directly to our defense and that most of them would

\(^{70}\) Ibid., 35.

\(^{71}\) Revenue Revision of 1941, 48, 56.
rather do it with their eyes open than do it through a general sales tax or through a multiplication of what we have known as ‘nuisance taxes.’”

Irritated by the president’s interference, Doughton rejected the request. But the Senate was more compliant, especially after Morgenthau caught up with Roosevelt’s recommendation. Echoing his boss, the secretary predicted that Americans would welcome a broader tax. “It would enable them to feel that they were participating personally and directly in the defense program,” he said. The Senate agreed, lowering exemptions to Roosevelt’s specifications. House lawmakers accepted the move, and experts predicted that five million new returns would arrive at the Treasury the next year. More than two million would show taxable income. The 1941 revenue act also hiked rates across the board; they now topped out at 77 percent on incomes over $5 million.

In March 1942, Morgenthau offered Congress a blueprint for tax hikes on an even larger, unprecedented scale. "Our task is more than the raising of a huge amount of new revenue," he told lawmakers. "It is to make the tax program an instrument of victory." Congress agreed, crafting a dramatic and far-reaching bill. Roosevelt later called the Revenue Act of 1942 the "greatest tax bill in American history," and the description may still be apt more than 60 years later. This sweeping measure is often credited with the creation of the modern U.S. tax regime. While it shares that distinction with several other pieces of wartime legislation, it was certainly the most important tax law of the war.

---

Initially, Morgenthau declined to recommend further exemption cuts, insisting that Americans making less than $15 a week were already heavily burdened (especially given the range of wartime excise taxes). Moreover, people exempt from the income tax were not a threat to price stability: “Their buying habits are governed strictly by the need of maintaining nutrition and health.” The secretary did, however, suggest a series of rate hikes for the income tax, ranging from less than one-half of one percent in the bottom bracket to more than 16 percent in the middle and upper reaches of the levy.74

Roosevelt’s newfound ardor for civic sacrifice and tax visibility, however, was not yet exhausted. On May 6, Morgenthau reluctantly suggested that exemptions be cut to $600 for individuals, $1,200 for couples, and $300 for dependents. The reductions would raise about $1 billion in new revenue and add about 7 million taxpayers to the rolls. Inflation worries lay behind some of these changes. “Prices were surging against their controls with the force of an angry sea,” recalled Randolph Paul.75 Leon Henderson told lawmakers than steep new taxes were absolutely vital if the country were to avoid ruinous inflation in the next year. Treasury experts agreed, urging Morgenthau to demand huge new hikes.76

But then Roosevelt caught his experts by surprise, yet again. In April, he proposed an apparently draconian measure to cap personal income. “[I]n time of this great national danger,” he declared on April 27, when all excess income should go to win the war, no American citizen

74 Revenue Revision of 1942, 297-300; Paul, Taxation in the United States.
75 Paul, Taxation in the United States, 300.
76 Ibid., 301; "Taxes Needed During Fiscal 1943 to Control Inflation," in Box 34; Inflation, Depression, Recovery; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1942).
ought to have a net income, after he has paid his taxes, of more than $25,000 a year.” The suggestion caused a brief firestorm of protest in the press, but it soon disappeared from serious discussion. Nonetheless, it demonstrated the president’s commitment to steep taxes on the rich. Roosevelt was still a fan of high-end progressivity, especially after wartime imperatives forced heavy taxes on the poor.78

In January 1942, barely a month after the attack on Pearl Harbor, Roosevelt had asked Congress for $50 billion to $60 billion in defense spending. By October, he had revised the figure upward to $76 billion. In March, Morgenthau asked Congress for $7.6 billion in new revenue to help support this massive spending request; later requests pushed the figure to almost $15 billion. Ultimately, however, Congress agreed to provide just under $7 billion: less than the March request and far less than the Administration had been seeking in the months since.

The act cut exemptions to $1,200, $500, and $350 for couples, individuals, and dependents, respectively. The “normal” tax increased from 4 percent to 6 percent, and surtaxes went up in every bracket. The first surtax rate more than doubled from 6 percent to 13 percent. But at the peak of the income scale, the changes were even more dramatic. The top rate climbed from 77 percent to 82 percent. Even more important, however, the brackets were readjusted to make this top rate applicable to a much larger percentage of national income. Under the 1941 law, the top rate kicked in at $5 million in annual income. But in 1942, the new, higher rate


78 The Treasury prepared a series of reports on the $25,000 income cap, ultimately, concluding that the tax would only affect about 11,000 Americans. Since it required a very large pre-tax income ($50,000 for an individual and $185,000 for a couple) to have a net income over $25,000, the tax was never a threat to most people. See Paul, Taxation in the United States, 301-02; "A Supertax on Individual Incomes above $25,000," in Box 54: Super Taxes; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1942).
started at just $200,000. A single taxpayer earning $1 million would see her tax climb from $655,139 to $809,995 (after accounting for the Victory tax). A married couple with the same income would see a similar hike. Clearly, rich Americans were being asked to pay handsomely.\footnote{Blakey and Blakey, "The Federal Revenue Act of 1942," 1071-72.}

*The Victory Tax*

As House and Senate lawmakers debated Treasury recommendations in 1942, they put their own stamp on the revenue bill. Perhaps their most important contribution was the “Victory Tax,” a short-lived income levy that nonetheless played a key role in the development of the modern tax regime.\footnote{For an excellent account of the Victory Tax, see Dennis J. Ventry, "The Victory Tax of 1942," Tax Notes 75 (1997).}

Lawmakers conceived of the Victory Tax principally as a revenue device, but its legislative odyssey hinged on its role in fighting inflation. The tax, according to its champions, would bolster the regular income levy by taxing Americans near the bottom of the income scale. Levied as a flat 5 percent tax on "Victory tax net income," it allowed an exemption of just $624. While regular income tax exemptions were also headed down in 1942, the Victory tax allowed fewer deductions from gross income, making it effective at much lower incomes. According to contemporary estimates, the tax would fall on 13 million new taxpayers.\footnote{Paul, *Taxation in the United States*, 319.}

The Victory tax was most notable, however, for an innovation it brought to the federal tax system: withholding. The tax was deducted directly from both salaries and wages, ensuring...
that taxpayers would remain current with their obligations. Current collection gave Treasury quick access to new revenue. It also helped restrain inflation by reducing discretionary income before it ever reached a consumer’s pocket.

A second innovation, even more fleeting than the Victory tax itself, was compulsory lending. The Treasury had been studying this means of inflation control for months, and the corporate excess profits tax actually included a refundable component similar to a compulsory loan. In general, compulsory loans were taxes collected during the war that were slated for repayment – either with or without interest – once peace returned. The Victory tax provided for a partial rebate of its 5 percent levy, effectively reducing the rate for most taxpayers to 3.75 percent for single persons and 3 percent for married couples. Treasury decided in 1944, however, that taxpayers should take the credit currently, thereby obviating a need for postwar refunds – and ending the experiment in compulsory lending before it even began.

In general, Treasury was never a fan of the Victory tax, complaining that it was complicated and regressive. But Congress liked the tax, largely because its raised a lot of money. And while the tax was regressive, it was consistent with a growing congressional fondness for shifting elements of the tax burden down the income scale. Rates for the personal income tax were already very high, reaching 82 percent in the 1942 bill. The excess profits tax was now levied at a flat 90 percent. There was, of course, still room for new taxes on the rich and

---

82 Ventry, "The Victory Tax of 1942." As Ventry explains, credits included: the lesser of 25 percent of the Victory tax or $500 for single people; 40 percent or $1,000 for married persons; and an additional 2 percent or $100 for dependents.

83 After its enactment, Treasury spent the rest of the war trying to fold the tax into the regular income tax. See "Integration of the Victory Tax with the Net Income Tax," in Box 51; Individual Taxpayers -- General; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1943).
corporate taxpayers. But lawmakers believed that taxes at the bottom of the income scale were a better choice. Not only would they raise important revenue, but they would also help thwart inflation.84

The Plan That Slogans Built

In 1943, Congress passed the Current Tax Payment Act, requiring taxpayers to stay current on their tax liability through a vast new withholding system applied to wages and salaries. The law represented a victory of sorts for the Treasury Department, where tax officials had been promoting current collection for years. Success, however, was a near thing, and it came at a steep price. Administration officials were forced to accept a tax forgiveness plan—championed by one of America's leading businessmen—that gave a windfall to rich taxpayers.

By 1943, withholding was already an old idea. During the Civil War, Congress had incorporated withholding into the fledgling—and fleeting—federal income tax. Later, during the early years of the 20th century, Congress again turned to withholding as it tried to craft a permanent federal income tax. (Withholding provisions were soon dropped from law, however, when administrative difficulties started to emerge.)

Policymakers of the World War II era had even more recent precedent for the advantages of withholding. In the years just before the United States entered the war, official Washington watched approvingly as new Social Security taxes were collected directly from wages and salaries. In impressive fashion, and without the use of computers, the Bureau of Old-Age

84 Committee on Finance, The Revenue Bill of 1942, 2 October 1942, 15-16.
Benefits managed to process over 312 million wage reports by mid-1940, posting more than 99 percent of them to 50 million employee accounts.85

Faced, then, with World War II's staggering revenue demands, tax officials saw withholding as a vital component of the new, vastly expanded federal income tax. Despite opposition from the Bureau of Internal Revenue (whose leaders never seemed to encounter an innovation they actually liked) Treasury officials began agitating for the reintroduction of withholding. The idea, they insisted, was administratively feasible and economically crucial.86

Treasury officials believed that withholding was necessary if the federal income tax were to function effectively as a mass tax. Before wartime revenue needs forced Congress to add millions of new taxpayers to the tax rolls, the income tax could be administered without withholding. But with so millions of new taxpayers, the system needed some mechanism of current collection, and withholding seemed the most promising. “The 39 million individual taxpayers required to pay income taxes of almost $10 billion for 1942 under the present law must be afforded a way of meeting their tax obligations with a maximum of convenience and a minimum of hardship,” Randolph Paul told Congress.87

Officials had already begun a massive public relations campaign to adviser new taxpayers of their fiscal responsibilities. In a multimedia campaign featuring posters, radio announcements, popular songs, and even a Donald Duck cartoon, Treasury drove home the filing requirements

86 “Collection at Source of the Individual Normal Income Tax,” in Box 54; Collection and Payment; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1941).
87 Randolph E. Paul, "Statement before the House Ways and Means Committee," in Box 54; Collection and Payment; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1943).
established by the new income tax laws. Officials stressed that millions of previously exempt Americans were now required to file annual returns.88

Treasury also argued that withholding would aid in the wartime battle against inflation. Withholding would withdraw more purchasing power more quickly from the economy, helping restrain the upward pressure on prices. Indeed, withholding could help transform the federal income tax into a potent tool for managing the nation's economy, allowing officials to regulate fiscal stimuli with enormous new flexibility.89

In November 1941, Morgenthau had presented a withholding plan to the House Ways and Means Committee, suggesting that new taxes should collected at the source from wages, salaries, interest, and dividends. A wary committee deferred his idea. The following March, Morgenthau tried again, and this time, lawmakers were more receptive. Ignoring protests from the commissioner of Internal Revenue – who insisted that the innovation would be expensive and labor intensive – legislators added withholding to the House version of the 1942 revenue bill. Senators were not yet ready, however, and the provision was dropped from the final bill.

But withholding wasn't dead yet. The final version of the Revenue Act of 1942 had included withholding for the new Victory tax. With that foot in the door, Treasury continued pressing the idea, arguing that Congress should extend this technique to the regular income tax. Roosevelt, too, lobbied for collection at the source, suggesting repeatedly in early 1943 that the income tax be put on a "pay-as-you-go" basis, a change that more or less required a withholding


89 Roy Blough, "The Individual Income Tax as a Method of Inflation Control," in Box 52; Individual Income Taxpayers; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1944).
mechanism.\textsuperscript{90} Such a system would keep taxpayers current, Roosevelt said, instead of allowing them to pay their taxes the year after they had incurred their tax liability.

Despite their rejection in 1942, current collection and withholding enjoyed considerable support on Capitol Hill. The BIR remained intransigent, complaining to Congress that withholding would be administratively burdensome and fundamentally unworkable given wartime personnel shortages. Randolph Paul, Treasury's general counsel and point man on tax matters, greeted the commissioner's statements unhappily. "It would be a masterpiece of understatement," he later wrote, "to say that these pronouncements, never submitted in advance by the Commissioner to the Treasury, were extremely embarrassing."\textsuperscript{91}

Despite his best efforts, BIR Commissioner Guy Helvering was unable to derail the drive for withholding. Only one serious obstacle remained: transition. Simply dropping withholding into the existing system would have required taxpayers to spend at least one year making "double" income tax payments. Under existing law, every taxpayer was expected to save enough money over the course of the year to cover his liability when it came due early in the next calendar year. If pay-as-you-go withholding were superimposed upon this system, then taxpayers would be further required to make simultaneous payments on their current liability.

Some observers saw nothing wrong with doubling up on tax payments during the transition year. The result was not really "double" taxation, they contended, since the lump sum payment and the current payments applied to income received during different years. If taxpayers

\textsuperscript{90} Franklin D. Roosevelt, "Annual Budget Message, 6 January 1943," (1943).

\textsuperscript{91} Randolph Evernghim Paul, Taxation for Prosperity (New York: Bobbs-Merrill, 1947), 331.
had been vigilant in their saving, they should have had enough money to pay the past year's liability, leaving them enough room in their paycheck for payments on the current year.

In fact, however, many taxpayers did not set aside enough money for their tax payment, making double payments a problematic cash-flow reality. Consequently, almost all policymakers agreed that some sort of transition relief would be necessary. Among the various possibilities floated by lawmakers and lobbyists, most involved the forgiveness of some or all tax liability for the year preceding the introduction of withholding.

Treasury tax officials acknowledged the transition problem, and they set to work on a solution. They refused to consider any blanket forgiveness of 1942 tax liability, arguing that it would disproportionately benefit the nation’s richest taxpayers.92 Similarly, they opposed "doubling up" since it, too, would favor rich taxpayers with savings adequate to pay additional taxes without hardship. Partial forgiveness or postponement of 1942 taxes were the only viable alternatives.

As Treasury officials mulled the options, a powerful challenge emerged from the private sector. For several years, Beardsley Ruml, chairman of the New York Federal Reserve Bank and treasurer of R.H. Macy and Company, had been campaigning for current collection, calling it a "simple" change of tax basis from the past year’s income to the current one. Effectively, Ruml was calling for a full year's tax forgiveness. Ruml pointed out that the Treasury would see no

92 Secretary Morgenthau told Roosevelt, "The plan presented about $64,000 to an individual with a net income of $100,000...for a man who had earned $2,000 only $140." See Jones, "Mass-Based Income Taxation: Creating a Taxpaying Culture, 1940-1952," 129.
reduction in cash flow. “[A]s far as the Treasury and income were concerned, things would move along just the same as time moves on under daylight saving,” Ruml contended. 93

Taxpayers immediately embraced the "Ruml plan." Popular speculation began to grow in late 1942 that Congress would pass legislation relieving all taxpayers of the need to pay their 1942 taxes. Treasury officials worried that people would be unprepared for the grim reality. House and Senate leaders issued public statement reminding citizens that payment was still expected by March 15, covering at least the first quarterly installment of tax against 1942 income.

Meanwhile, Roosevelt rejected the Ruml plan out of hand: "I cannot acquiesce in the elimination of a whole year's tax burden on the upper income groups during a war period when I must call for an increase in taxes and savings from the mass of our people," he declared. Treasury officials echoed Roosevelt's point, arguing that forgiveness "would bestow the greatest benefit on those best able to pay and the smallest benefit on those least able to pay."94

Congress and the administration agreed that resolution of the withholding issue was vital to the nation's economic health. Treasury officials restated their support for current collection, stressing the administrative need for such a system. But they also stressed their opposition to a full year's tax forgiveness. While acknowledging that revenue flows would be uninterrupted, they insisted that forgiveness was unfair. In its stead, they proposed a scheme of deferred payments, under which taxpayers would be required to pay at least a large percentage of their 1942 tax liability but would be allowed several years in which to do so.

93 Paul, Taxation in the United States, 329.
94 Brownlee, Federal Taxation in America, 92, 116.
Ruml weighed in against the Treasury plan, arguing that the government could forgive a full year's taxes without unduly burdening itself. The asset loss incurred by tax forgiveness would only be evident when examining the Treasury's position on Judgment Day. At that point, millions of Americans would die owing the government money. “[T]hese would be bad debts in any case,” Ruml observed wryly, so the government had nothing to lose. While Ruml's original plan did not include provisions for withholding, he now added it to the list of reasons for supporting his plan. Collection at source, he declared, was fully consistent with his forgiveness plan.

The House Ways and Means Committee rejected the bulk of Ruml's argument, with Democratic members insisting that forgiveness was little more than a windfall for the rich. The full House, however, was more sympathetic, and the Ruml plan continued to surface in floor votes. The final bill emerging from the House embodied a modified forgiveness plan directing most benefits to taxpayers in the first income bracket. The Senate Finance Committee, for its part, adopted the Ruml plan and its full-year forgiveness.

As the conference began, House and Senate leaders remained firmly at odds, the former digging in their heels against "excessive" forgiveness. Additionally, the issue had taken a distinctly partisan tone, with Republicans generally favoring the Ruml plan and Democrats denouncing it. Roosevelt wrote congressional leaders to emphasize his strong support for pay-as-you-go legislation and renew his strenuous objection to full-year forgiveness. Excessive forgiveness, he asserted, "would result in a highly inequitable distribution of the cost of the war...

---

Paul, Taxation in the United States, 335.
and in an unjust and discriminatory enrichment of thousands of taxpayers in the upper income groups.”

In fact, Roosevelt was well aware of those taxpayers who stood to benefit. In March, he had asked Treasury to draw up a list of taxpayers likely to reap the greatest benefits; “no names, of course,” he added. The Treasury responded with a memo listing the 100 largest federal taxpayers – by name – along with their projected savings under the Ruml plan. The memo read like a roster of Roosevelt's staunchest opponents, although it also included a variety of celebrities and even the odd Democrat.

After lengthy delay, the conference deadlock ended when Ways and Means Chair Robert Doughton changed his vote, allowing a compromise bill to emerge. The final version of the bill included a substantial tax forgiveness – less than Ruml had sought, but considerably more than Treasury wanted. The Current Tax Payment Act of 1943 provided for current payment of all individual income tax liabilities and the cancellation of 75 percent of one year's existing taxes (the lower of either the 1942 or 1943 tax liability). Unforgiven liabilities were payable in two installments, one on March 15, 1944, and the other on March 15, 1945.

The introduction of withholding and the debate over the Ruml plan are significant for two reasons. First, they reveal the extent to which Roosevelt and his administration were willing to fight for tax fairness. The president and his tax specialists consistently opposed complete forgiveness, complaining that it would provide a windfall to the nation's wealthy. FDR’s request for a list of rich taxpayers was consistent with his longstanding tendency to personalize abstract

---

issues of tax justice. He always viewed tax issues in terms of winners and losers, and more often than not, he was willing to identify them. (Although in this case, discretion seems to have won out.)

Second, withholding changed the income tax forever. It made the levy more responsive and flexible, both reflecting and facilitating its conversion into a powerful tool of macroeconomic regulation. Moreover, as one legal historian has pointed out, it helped create a taxpaying culture, getting Americans comfortable with regular deductions from their paychecks. No small feat in an era when such deductions were all but unknown.98

“Relief Not for the Needy but for the Greedy”

In late 1943, Henry Morgenthau asked Congress for another tax increase to provide $10.4 billion in new revenue. That figure included $6.5 billion from the individual income tax, $400 million in estate and gift taxes, $1 billion in new corporate taxes, and $2.5 billion in excise tax increases.99 In an effort to simplify the tax system, the administration also asked for repeal of the Victory Tax, a flat-rate income levy imposed on top of the normal income tax. The tax applied to many Americans not otherwise subject to income taxation.

Morgenthau framed his appeal, in large part, as an anti-inflation effort. The new act, he argued, must continue the battle against inflation. "[N]othing in the economic field can interfere

---

98 Jones, "Mass-Based Income Taxation: Creating a Taxpaying Culture, 1940-1952."

99 The administration’s proposed estate tax changes would have boosted rates, pushing them to 80 percent on sums more than $1.5 million. It would also have lowered the estate tax exemption to $40,000. Supporters insisted that the levy made only a small contribution to total revenue, and that rates in lower and middle brackets were still moderate. Indeed, while estate tax increases had been a fixture of wartime tax reform, they played a distinctly secondary role in the federal tax system; between 1940 and 1945, their contribution to total receipts dropped from 6.7 percent to just 1.3 percent. See Roy Blough, "Postwar Tax Structure," in Box 64; Postwar Planning; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD. (1944).
with the war effort as much as an uncontrolled rise in prices," he told the House Ways and
Means Committee, reiterating his comments from the year before. "An inflationary price rise is a
source of grave social injustice."

Congressional taxwriters didn’t want to hear such language. Indeed, Morgenthau’s
package got a chilly reception. Lawmakers fell on the plans "like Caesar's assassins," recalled
one observer. Increases for income and estate taxes took much of the heat, but so, too, did the
suggestion to repeal the Victory Tax. Repeal would have removed nine million low-income
taxpayers from the rolls, but congressional leaders found little to like in that prospect. “Almost
before the Secretary had finished reading his prepared statement,” reported Time magazine. “the
U.S. Treasury's design for extracting another $10,500,000,000.00 from the U.S. pocketbook was
mackerel-dead.”

Various administration officials echoed Morgenthau's call for higher taxes. Federal;
Reserve Chairman Marriner Eccles asked for even bigger numbers: $13.8 billion, including $4
billion in taxes that would be refundable after the war. Ways and Means Chairman Bob
Doughton, D-N.C., dismissed the higher figure out of hand. "Amazing, fantastic, and visionary. I
don't like it at all. If possible, it is worse than the Treasury program," he declared.

Ways and Means went to work on its own bill, which the House passed on November
24. The legislation bore little resemblance to the Treasury plan. Committee members argued that
the need for new taxes had been diminished by revenue already in the pipeline. Coupled with

100 Committee on Ways and Means, Revenue Revision of 1943, 78th Congress, 1st Session, 4-8, 11-16, 18-
20 1943, 4.
101 Paul, Taxation in the United States, 356; "The High Cost of Morgenthau."
102 Paul, Taxation in the United States, 358.
promised reductions in nonmilitary sending, existing taxes would be nearly adequate. As for inflation, the panel declared that it could be controlled through discretionary spending cuts, effective price controls, rationing, and wage limits. Higher taxes were unnecessary.  

The Ways and Means Committee specifically opposed corporate tax increases. “It is vitally important that our corporations be kept in sound financial condition so that they may be able to convert to peacetime production and provide employment for men leaving the armed forces after the war,” the lawmakers declared. In fact, complaints about corporate tax burdens were on the upswing by 1943, and by 1944, cuts seemed almost inevitable.

New revenue in the Ways and Means bill came from several sources, including repeal of the earned income credit and a new minimum tax for the Victory Tax; both provisions targeted low-income taxpayers. Excises provided another $1.2 billion, with much of the revenue coming from a 50 percent liquor tax hike. The bill reduced the invested-capital credit for large corporations, but increased the exemption for the excess profits tax from $5,000 to $10,000. Rates for the excess profits tax went from 90 to 95 percent. Postal rate hikes added a little more money, bringing the bill's total to about $2 billion – less than a fifth of the Administration's request.

The House bill got a bad reception, both from Roosevelt and the press. Editorials across the country denounced the bill as inadequate. Meanwhile, Morgenthau tried to make up lost ground by working closely with the Senate Finance Committee. While eager to placate
conservatives uneasy with the Administration's penchant for steep estate and income taxes, he nonetheless urged lawmakers to go easy on the little guy. The bill coming from the House already drew more than 50 percent of its revenue from people making less than $5,000, he pointed out.

Senators were unmoved. People earning less than $5,000 might be paying half the new taxes, but they also received four-fifths of the nation’s income, critics pointed out. If the Administration was serious about fighting inflation, then low and middle income Americans would have to shoulder a bigger share of the tax burden. “The principal Congressional objection to the Treasury plan,” reported the *New York Times*, “has centered on the Administration’s apparent unwillingness to increase taxes on low incomes.” The Finance Committee endorsed the general approach of the House bill and opposed most additional increases. The panel tinkered with the Victory Tax, and put off a long-planned increase in the Social Security payroll tax. The final bill emerging from the conference committee closely resembled the Senate package, and Congress sent it to the President in February 1944.

On February 22, Roosevelt vetoed the bill. In explaining his decision, he cited its inadequate revenue yield. The most serious problem, however, was its tendency to compromise sound tax policy for the sake of political expediency. Riddled with numerous loopholes, most directed at business, the law was a travesty. "In this respect," the president complained, "it is not a tax bill but a tax relief bill providing relief not for the needy but for the greedy."

---

Roosevelt objected to the elimination of planned increases in the Social Security tax, a move that ensured a substantial cut in anticipated revenue. But he particularly scored a variety of provisions offering "indefensible special privileges to favored groups." Those privileges set a bad precedent, even threatening the viability of the tax system; by degrading fairness, they undermined the political consensus so fundamental to wartime tax policy. Roosevelt noted that some of his advisers wanted him to sign the bill, arguing that having asked for a loaf of bread, he should be content with a small piece of crust. "I might have done so if I had not noted that the small piece of crust contained so many extraneous and inedible materials," Roosevelt said.

Roosevelt's caustic veto message enraged legislators. Ways and Means Chairman Doughton declared his intention to seek an immediate override; 19 of the committee’s 25 members signed a public letter declaring the Treasury’s original plan a threat to the nation’s economy. Finance Committee Chairman Walter George, D-Ga., opined on the overwhelming burden of wartime taxation. The day after the veto, Senate majority Leader Alben Barkley, D-Ky., resigned his leadership post in protest over the president’s action. The president’s veto, he declared on the Senate floor, was “a calculated and deliberate assault upon the legislative integrity of every member of Congress.” Congress had a simple choice: override the veto or surrender the legislature’s role in the tax policy process.107

Roosevelt tried to calm things down, asking Barkley to rethink his decision. He also pointed out that Barkley had been read large sections of the veto message before it was released, making the senator’s public declaration of surprise something of a surprise in itself. But Barkley

107 Paul, Taxation in the United States, 373.
followed through with his resignation, only to have his Senate colleagues immediately re-elect him to the same post. The charade made headlines, underscoring Roosevelt’s predicament.\textsuperscript{108}

Within a week of FDR’s veto, Congress voted to override it, the measure passing with a large majority in both houses. As enacted, the 1943 Revenue Act failed to satisfy the twin imperatives of World War II tax policy: revenue raising and inflation control.\textsuperscript{109} It succeeded, however, in allowing legislators to duck the responsibilities of wartime policymaking, and it even made room for a few special favors.

The Revenue Act of 1943 was a sign of things to come. The congressional penchant for handing out tax preferences would soon become a hallmark of postwar tax legislation. And while presidents of both parties – as well as a few key legislators – would continue the fight for the fisc, it was a Sisyphean struggle. The modern income tax would be riddled with loopholes. Indeed, the steep progressive rate structure made this almost inevitable, raising the stakes for almost all taxpayers and making every tax preference more valuable.\textsuperscript{110}

\textsuperscript{108} Ibid.

\textsuperscript{109} Among its various major provisions, the Revenue Act of 1943: left individual income tax rates, as well as exemptions, unchanged; left estate tax rates and exemption unchanged; cut the Victory tax rate from 3 percent to 5 percent, but repealed credits allowed under the levy; raised the excess profits rate from 90 percent to 95 percent but also raised the exemption from $5,000 to $10,000; raised a range of excise taxes, totaling more than $1 billion in new revenue. For a summary, see Ibid., 375-78.

Conclusion

In his noted study of federal taxation, political scientist John Witte described the income tax revisions of World War II as “a series of incremental adjustments to existing policies.”¹¹¹ Perhaps. But that’s like telling the man at the end of a gangplank that his next step is just like any other: some increments are more important than others. The exemption cuts of World War II totaled just $500 for individuals, but they transformed the nature of the American state and society. Similarly, the rate changes, while often simply a matter of adjusting numbers in a table, were enormously important to the taxpayers who suddenly found themselves with marginal tax rates over 90 percent.

In several respects, the wartime tax regime broke with the history of New Deal revenue reform. The broad-based income tax forced millions of middle class Americans onto the tax rolls for the first time; no strangers to federal taxes on consumption, they were unaccustomed to paying direct taxes like the income levy. The war gave this erstwhile class tax a new suit of clothes. Almost overnight, one legal scholar observed, the income tax “changed its morning coat for overalls.”¹¹²

This change was a victory for advocates of low-end progressivity, who had long maintained that income taxes should take the place of consumption taxes in the federal tax system. Wartime revenue needs made both kinds of taxes necessary, but the decision to avoid a new federal sales tax represented a major achievement for income tax supporters. Absent strong


Administration leadership, Congress would almost certainly have opted for some sort of sales tax, despite the objections of organized labor. Roosevelt’s embrace of low-end progressivity – and the mass income tax in particular – proved decisive.

At the same time, however, FDR did not abandon his preference for heavy, soak-the-rich taxation. He demanded steep rates for both corporate and individual taxes, firm in his conviction that the rich must shoulder a heavy burden. Indeed, the wartime revenue regime represented a blend of high-end and low-end progressivity. In the decades to come, its resilient political balance – trading steep rates on the rich for broad taxes on the poor – would offer a solid foundation for federal finance.


———. "The Individual Income Tax as a Method of Inflation Control." In Box 52; Individual Income Taxpayers; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD., 1944.

———. "Postwar Tax Structure." In Box 64; Postwar Planning; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD., 1944.

"Broader Tax Base Studied by House Unit." Washington Post, Aug 8, 1939, 11.


"Collection at Source of the Individual Normal Income Tax." In Box 54; Collection and Payment; Records of the Office of Tax Analysis/Division of Tax Research; General
"President Suggests Tax Base Be Wider." Washington Post, June 17, 1939, 1.
"Proposal for a 'Consumption Expenditure Tax'." In Box 6; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO., 1942.
"Retailers Propose 5% Sales Tax to Raise $3 Billion." Wall Street Journal, 1 April 1942, 5.
Committee on Finance. Revenue Act of 1940, 77th Congress, 1st Session, 12-14 June 1940.
Committee on Ways and Means. The Revenue Bill of 1940, 76th Congress, 3rd Session, 10 June 1940.
Committee on Finance. The Revenue Bill of 1942, 2 October 1942.
Committee on Ways and Means. Revenue Revision of 1941, 77th Congress, 1st Session, 24, 28-30 April, 1-2,5-9 May 1941.
Committee on Ways and Means. Revenue Revision of 1943, 78th Congress, 1st Session, 4-8, 11-16, 18-20 1943.
———. "Letter to Congress on Tax Bills." 1943.
"Sales Tax Gains Favor with House Committee." Washington Post, Jan 9, 1942, 1.
"A Scheme for a Progressive Sales Tax." In Box 44; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO., 1942.

62
———. "Plans for Additional Revenue." In Box 43: Methods of Raising Additional Revenue; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD., 1938.
"The Spendings Tax." In Box 6; Papers of Roy Blough; Harry S. Truman Presidential Library, Independence, MO, 1942.
"A Supertax on Individual Incomes above $25,000." In Box 54; Super Taxes; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD., 1942.
"Taxes Needed During Fiscal 1943 to Control Inflation." In Box 34; Inflation, Depression, Recovery; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD., 1942.
"Treasury Anti-Inflation Tax Proposals." In Box 34; Inflation, Depression, Recovery; Records of the Office of Tax Analysis/Division of Tax Research; General Records of the Department of the Treasury, Record Group 56; National Archives, College Park, MD., 1942.


THE PROBLEM OF CHARITABLE INSIDER TRADING

Darryll K. Jones

I. Introduction

It is, by now, axiomatic that markets function best when participants enjoy unhindered access to pricing information. American law strives to maintain equal rights to information because it is assumed that when both sides to any transaction have accurate information everyone will be satisfied with the outcome; society will be satisfied because goods and services flow to their highest and best use. For this reason, the law withholds respect from transactions characterized by information imbalances, at least when the imbalance is caused by the actions of one of the participants. In fact, the law actively disrespects transactions consummated such under such circumstances. The self-interests upon which markets best function simply cannot be exercised in the presence of information imbalances.

Equal information access is not, by itself, sufficient to ensure the proper functioning of the relevant market. Threats persist even when all is known by all concerned if one party is under duress or the control of another, or when one party assumes the responsibility of putting freely available information to use on behalf of

Markets are efficient when prices accurately reflect all available information regarding the assets traded in the market. Attaining efficient pricing is crucial for achieving efficient allocation of resources in the economy. Among other things, efficient pricing is important for the market for corporate control, for monitoring and controlling the management agency problem, for the allocation of resources through initial public offerings (“IPOs”) and secondary offerings, for keeping high liquidity in the market, and for other transactions in the economy that rely on market prices. Markets are liquid when traders can execute transactions speedily. The more liquid a market, the faster a trading order is executed. Liquid markets benefit the economy as they reduce the cost of transacting and the risk associated with investment.


2 This observation is not only true with respect to financial transactions. It is equally applicable to other sorts of transactions, including the making of prenuptial agreements, the conduct of civil and criminal litigation, and the making of contracts, all of which to are respected only when certain information disclosures of one sort or another are made.

3 The “disclose or abstain” rule, for example, is meant to ensure the integrity of capital markets by allowing insiders to trade only when both sides to a stock transaction have access to the same information. See 17 C.F.R. 240.10b-5 (1981).
The recentl y adopted Sarbanes -Oxley Act is large ly, if not excl usive ly conc erned wit h fi xing f lawed 
proces ses so t hat informat ion i s made a vail able for unhi ndered use by  all  part ies.  Sar banes -Oxl ey Act  of 

Critical analysis an d sum mary o f Sarb anes-O xley is ubiquitou s alread y.

See , e.g., Ne ils Scha uma nn, The 
Sarbanes Oxley Act :  A Bir d’s- Eye View, 30 

5 The modern a pproach to con tracts made unde r ci rcums tanc es s ugges ts information imba lance or 
barrier s to use caused by one party is to make the contract voidable as a matter of right granted to the other 
party.  See, e.g., Restatement (Second) of Contracts, §7, Comment b (1981) (relating to grounds of 
contract avoidance).

6 The recently adopted Sarbanes-Oxley Act is largely, if not exclusively concerned with fixing flawed 
processes so that information is made available for unhindered use by all parties. Sarbanes-Oxley Act of 
Critical analysis and summary of Sarbanes-Oxley is ubiquitous already. See, e.g., Neils Schaumann, The 


8 Id. At ___. (describing the “veil of ignorance”).

9 Most recently, however, nonprofit law has begun to offer incentives to insiders who utilize demonstrably 
correct procedures; the statutory offer, though, is accompanied by a statement eschewing any negative 
presumption from the failure to follow the process. See Treas. Reg. 53.4958-6 (2002) (describing a process 
which, if followed, leads to a rebuttable presumption that the transaction will receive the benefit of legal 
recognition and providing that no negative inference arises from the failure to follow the described 
process).
over with respect to which she occupies the status of insider.\textsuperscript{10} It is, instead, a noun phrase though the resulting harm is commensurate with the harm from the more familiar sort of insider trading. The market harm results not from the denial of information but from the organization’s inability to use information unhindered by the insider’s conflicting loyalties. The genesis is in the failure of process. The law’s reaction to flawed process is quite nearly visceral when it relates to conflicts of interest in for-profit transactions\textsuperscript{11} but, as this article shows, the law reacts with ambivalence, at best, with regard to charitable insider trading characterized by flawed process.

The harm referred to in this article is primarily to the reputational aspect of nonprofit markets. Markets thrive or falter on reputation.\textsuperscript{12} For-profit insider trading, for example, might have relatively slight or insignificant direct harm. The harm might even appear temporary and limited to the single organization involved in the transaction. In isolation, or in combination with a relatively few more instances, though, for-profit insider trading can have devastating effects on the market’s reputation.\textsuperscript{13} Investor confidence suffers and the fear is that investors will withdraw the capital upon which the

---

\textsuperscript{10} In this regard charitable insider trading is more a function of status rather than knowledge. For-profit insider trading refers to the act of trading in securities by persons in possession of nonpublic information. See Jesse M. Fried, \textit{Insider Abstention}, 113 \textit{Yale L. J.} 455, 458-59 (2003) (distinguishing between three types of insiders based on the relationship giving rise to their knowledge of nonpublic information). Nonprofit insiders are simply those with managerial control, regardless of their actual knowledge. United Cancer Council v. Commissioner, 165 F.3d 1173, 1176 (7th Cir., 1999). See also, I.R.C. 4958(f) (2006) (describing “disqualified persons” – known colloquially as “insiders” – in part as any person “in a position to exercise substantial influence over the affairs of the organization.”).

\textsuperscript{11} See, e.g., \textit{Revised Model Business Corp. Act} §8.61 (1984) (regarding director liability for conflict of interest if certain processes designed to ensure complete disclosure and unhindered use of information).

\textsuperscript{12} Consider the following insightful and pragmatic discussion of the importance of reputation to business markets:

Financial institutions and regulators respond to the weakening trust in the markets by assuring investors that their suspicions are unfounded. One can analogize the trustworthiness of the system to reputation. To reap future benefits from the financial markets, financial institutions and Congress invest in building up the markets' reputation for trustworthiness. But when times are good, Congress and the industry overconsume the fruit of a good reputation and deplete its value. When events lead investors to question the system's trustworthiness, Congress, the issuers and the industry start to repair the damage and invest in market reputation again. When investors' trust is restored, industry pressures increase to reduce the cost of capital formation. Investors do not seem to mind. Regulation is relaxed. And when prices fall, the cycles start all over again.

Congress, the issuers and the industry should learn what private sector corporations know so well in the context of their businesses. Deplete your reputation with shoddy goods and you are bound for lean years of meager profits and high investment to shore up your reputation. The benefits from the depletion may or may not be worth it. Congress, the issuers and the industry knew this lesson after tightening regulation in the 1930s and 1970s.

market survives. Those effects, in turn, interfere or prevent the market’s ability to be the provider of desired goods and services.

The market interference or prevention conclusion applies as well to the non-profit sector. Rumor and scandal provoked by charitable insider trading does as much harm to nonprofit markets as insider trading and other manifestations of conflicted interests have in for-profit markets.\(^{14}\) Direct stakeholders in the nonprofit sector rightly view rumor and scandal as a sort kryptonite, regardless of how infrequent those incidences may be.\(^{15}\) The most feared market impact, of course, is that donors will withhold their nonprofit sector funding. That is, donors will not patronize nonprofit organizations because they will not trust that their patronage benefits those on whose behalf the organization purportedly operates.\(^{16}\) While existing evidence is not entirely empirical, there is widespread consensus that charitable contributions invariably decrease after widely publicized scandal and rumor.\(^{17}\) Donors in non-profit markets are sufficiently analogous to consumers in for-profit markets that it is not surprising that informational misuses in non-profit markets harm that sector just as such misuses harm the for-profit sector.

As noted earlier, the law protects the for-profit market with respect to information imbalances and conflicted transactions in many ways. With respect to the nonprofit market, though, the law is hardly sufficient. Except with regard to private foundations, nonprofit law almost encourages transactions involving inherent informational imbalances or use barriers.\(^{18}\) Where for-profit laws manifest suspicion and hostility from the imminence of information abuse, nonprofit laws withhold suspicion and hostility until the actualization of abuse, and then usually only when that abuse is extreme.\(^{19}\)


\(^{15}\) See Robert O. Bothwell, Trends in Self-Regulation and Transparency of Nonprofits in the U.S., 2 The Int’l J. of Not-For-Profit Law, available online at http://www.icnl.org/journal/vol2iss3/Arn_bothwell.htm (last visited March 28, 2006). (regarding the fear among charitable leaders that donors will “lose confidence in the good of charitable organizations”).

\(^{16}\) Leslie G. Espinoza, Straining The Quality of Mercy: Abandoning The Quest for Informed Charitable Giving, 64 S. Cal. L. Rev. 605, 664 (1991) (regarding the effects of a loss of donor confidence brought about by the suspicion that organizations do not use donations for their intended purpose).


\(^{18}\) The law is arguably excessively vigilant with regard to organizations designated as “private foundations.” For example, present law imposes an unavoidable excise tax on nearly all transactions between insiders (and all their relatives and business associates) and the charitable organization. See IRC 4941 (2006).

\(^{19}\) IRC 501(c)(3), inter alia, prohibits insiders from “siphon[ing] its earnings to its founder, or the members of its board, or their families, or anyone else fairly to be described as an insider, that is, as the equivalent of an owner or manager. United Cancer Council v. Commissioner, 165 F.3d 1173, 1177 (7th Cir. 1999). IRC 501(c)(3) (2006). A more recently enacted provision impose excise taxes directly on insiders who engage in such behavior. IRC 4958 (2006). Neither provision, nor any regulation or other authority addresses processes that would reduce the likelihood of the prohibited harm.
There are reasons, largely mythical perhaps, why this is so. They have to do with the romanticized notion that nonprofit insiders invariably occupy the dual role of benefactor as well as service provider.\textsuperscript{20} As will be shown, in most other respects the law has long since abandoned the notion that insiders are as much benefactors, if at all, as they are plain old self-interested employees. The myth of insider as well-intentioned benefactor helps explain why the law is not at all suspicious when insiders trade with the charitable organization over which she exercises managerial authority. The myth makes it easy to discount, and then not adequately regulate against, the imminent harm from charitable insider trading.

It would be incorrect, though, to conclude that the law is entirely ambivalent to the harm resulting from charitable insider trading. In this regard, incidentally, I define “the law” as the federal income tax provisions pertaining to the grant of tax exemption.\textsuperscript{21} Although state laws purport to regulate nonprofit organizations, they are ignored in this article primarily because those laws are generally ignored in the real world.\textsuperscript{22} State law simply has no practical relevance in regulating the problem with which this article is concerned. Federal tax law has expressed formal interest, at least, in the problem since its original enactment.\textsuperscript{23} The first federal tax law exempted charitable organizations from tax only on the condition that insiders not use their managerial authority over the organization to reap individual profit.\textsuperscript{24} As explained below, this is a result-oriented prohibition rather than a process-oriented prohibition of the sort more often imposed in for-profit markets. My quarrel is therefore not that the law is entirely ambivalent but that it manifests insufficient suspicion and hostility to charitable insider trading; the insufficiency prevents a proactive approach with regard to protecting the nonprofit market reputation. The purpose of this article, frankly, is to assert the need for a justifiably higher level of suspicion and hostility with respect to charitable insider trading.

Present law sets aside even justifiable suspicion and hostility, and reacts instead solely on the occasion of actualized harm. A very strong argument can be made that the law should maintain a certain suspicion and hostility before the harm occurs. If market sector reputation is important, waiting until harm actually occurs rather than reacting defensively to the whiff of harm is “too little, too late.” A sufficiently defensive reaction to imminent harm would require a bright line rule completely prohibiting insiders from

\begin{itemize}
\item \textsuperscript{20} Nonprofit board members are most often invariably benefactors to the extent they serve without compensation. See Wendy K. Szymanski, An Allegory of Good (and Bad) Governance: Applying the Sarbanes-Oxley Act to Nonprofit Organizations, 2003 Utah L. Rev. 1303, 1316-17. It is also frequently the case that board members become or are expected to be donors, as well.
\item \textsuperscript{21} See infra note ____.
\item \textsuperscript{22} Scandals in the for-profit world, coupled with state and federal legislation in response thereto, appears to be provoking more state law enforcement of fiduciary responsibilities in nonprofit organizations. See Wendy K. Szymanski, supra note ____ at 1304-05 (regarding state efforts to adapt and apply Sarbanes-Oxley to nonprofits).
\item \textsuperscript{23} IRC 501(c)(3), containing the prohibition against insider “siphoning,” has existed largely unchanged since its first appearance. See Tariff Act of 1909, Ch. 6, §38, 36 Stst. 11, 113 (1909) (containing the prohibition against private inurement).
\item \textsuperscript{24} Id.
\end{itemize}
transacting with charitable organizations, except with regard to their status as employee.\textsuperscript{25} A rule such as this has never been adopted most likely because the myth of insider as benefactor has not been entirely debunked. Such a rule would be process--, rather than result-oriented. It would allow an immediate response to a process – charitable insider trading – so pregnant with harmful consequence as to make the result irrelevant. In this article, though, I assume the present result oriented approach will continue and take issue with the universe of results about which present law eventually, if not belatedly, responds.

If asked, donors would likely identify overpayments to nonprofit managers as the transaction that most affects their willingness to participate in the nonprofit market. There seems nothing more sensational, in a negative sense, than ostensible “do-gooders” who are caught pilfering and profiteering from their organization at the expense of charitable beneficiaries. The legal phrase used to describe the event is “private inurement”\textsuperscript{26} or, more recently, “excess benefit transactions.”\textsuperscript{27} Private inurement and excess benefit are labels appropriate to the myriad ways by which organization managers – insiders – “siphon” off a charitable organization’s assets for their own personal use and without providing quid pro quo.\textsuperscript{28} It is a result-oriented rule because there is neither violation nor sanction when an insider transacts with the organization unless the result is actually harmful. As such, the rule ignores the great potential for harm resulting from the insider’s ability to control the organization’s knowledge and use of relevant information. Thus, private inurement does not occur when an insider is paid a market salary for her labor on behalf of the organization. Private inurement occurs, for example, when an organization buys supplies from an insider’s private business, but only if the organization pays too much for those supplies. The law is entirely unconcerned with the process that may have led to the organization’s decision to purchase supplies from an insider’s for-profit entity. It should not be surprising, though, that donors might be concerned that their patronage is being used to provide a customer base for the insider’s private business.

In a 2000 article, I defined and categorized three generic transactions that result in private inurement.\textsuperscript{29} The first type of transaction is labeled “strict accounting private inurement.” Strict accounting private inurement means only that the organization has paid too much or charged too little for necessary goods or services.\textsuperscript{30} An insider, for example, might cause a charitable organization over which she has control to pay an excessive salary. “Excessive” is determined, of course, by reference to the market. The second type of transaction is labeled “incorporated pocketbook private inurement” and

\textsuperscript{25} Indeed, this is the approach taken with respect to private foundations. See infra note ____.
\textsuperscript{26} IRC 501(c)(3) grants exemption for certain organizations provided “no part of the net earnings of which inures to the benefit of any private shareholder or individual.” IRC 501(c)(3) (2006).
\textsuperscript{27} IRC 4958, enacted in 1996, imposes personal liability on insiders who engage in private inurement but refer to the transactions as “excess benefit transactions.” IRC 4958 (2006).
\textsuperscript{28} For an exhaustive discussion of the ways in which private inurement and excess benefit occurs see, Darryll K. Jones, The Scintilla of Individual Profit: In Search of Private Inurement and Excess Benefit, 19 VA. TAX REV. 575 (2000).
\textsuperscript{29} Id. at 590-639.
\textsuperscript{30} Id. at 595, n. 69 (citing cases exemplifying strict accounting private inurement).
refers to an insider’s use of an organization’s assets for her personal use and without any benefit to the organization. With incorporated pocketbook private inurement, the organization becomes the insider’s private bank from which she makes withdrawals as her personal needs and desires dictate. When an insider uses an organization's assets to fund lavish vacations, under the guise of “business,” for example, she is engaging in incorporated pocketbook private inurement. More frequently, though, incorporated pocketbook is comprised of many smaller instances of personal expenditures that aggregate to such an extent over time that the insider’s funded consumption becomes grossly conspicuous. The result is too obvious to ignore. Unfortunately, these sorts of transactions have taken on the appearance of ubiquity and, as a result, endanger the reputation of charitable markets.

It is worth stating one final time that the law prohibits these results, just as it does so in for profit markets. Unlike for-profit markets, the law imposes no sanctions or rules in the nonprofit market with respect to circumstances or processes that should justify suspicion or hostility before the actualization of harmful consequences. There are, in fact, no mandates or prohibitions concerning the circumstances that give rise to the inevitability. Since the inevitability of harm is not a cause for reaction, and the sole cause for reaction is the actualization of harm, there is one other type of private inurement that escapes sanction more often than not because the actual harm is even more insidious. The third type of transaction, first proposed in my earlier article, is referred to as “joint venture private inurement.” Joint venture private inurement recognizes that an exempt organization is, by proxy, a consumer on behalf of all of its beneficiaries. The consumptive power derives from donations and other income provided to the organization and, indirectly, to the organization’s beneficiaries. The greater the number of beneficiaries, the greater the organization’s wealth maximizing effect is for those who want to serve those consumers. That consumer power, obviously, should be expended in a manner most advantageous to the beneficiaries and the harm when this is not the case is most acute for beneficiaries with the highest level of dependence.

My earlier article advocated applying the private inurement prohibition to transactions in which an insider instead monopolizes an organization’s invariable consumer power for her own benefit rather than for beneficiaries’ highest benefit. In essence, joint venture private inurement occurs when an insider causes an organization to grant a franchise — comprised of the organization’s consumers — to the insider via transactions solely with the insider’s profit making apparatus, even if those transactions are at market rates. “The violation occurs because the operations of the tax exempt entity and an insider-controlled entity are so closely related that the insider, by virtue of her interest in the taxable entity, financially benefits from the exempt entity’s invariable consumer power.” In terms of the theory espoused in this article, the organization’s inability to use its consumer power unimpeded by the insider’s interest invariably results in harm to the nonprofit market. The harm proceeds from the organization’s inability to

---

31 Id. at 610, 611 n. 113 (citing cases exemplifying incorporated pocketbook private inurement).
32 Id. at 620, 623 n. 138 (citing cases in support of joint venture private inurement).
33 Id. at 620.
use information about the market to best serve its beneficiaries. Whether that is the ultimate result, though, is immaterial. The harm is that the process unnecessarily makes the result possible.

This article proceeds from my first article and is more directly provoked by two recent events. The earliest event involves the Internal Revenue Service’s reaction to my assertion of a “joint venture private inurement” theory. On September 9, 2005 the Internal Revenue Service (“the Service”) issued proposed regulations that apparently accept the notion of joint venture private inurement.34 Unfortunately, the Service mislabels the transaction as a private benefit transaction rather than a private inurement or excess benefit transaction. The former label applies even less suspicion than that applied to private inurement or excess benefit transactions. In short, the proposal allows incremental harmful results so long as the insider maintains a credible façade indicating that the organization is operating substantially for its charitable beneficiaries. Here again, the rule depends too greatly on results, and then only in very extreme cases.

The second event provides a vivid example and case-study of the type of harm discussed in this article. It also serves to demonstrate the ineffectiveness of present law and the Service’ proposed method of implementing the joint venture private inurement prohibition. When Coretta Scott King died in early 2006, her funeral was held at New Birth Missionary Baptist Church, whose Pastor is Bishop Eddie Long. Quite understandably, the Church became the subject of significant press coverage. The Atlanta Journal-Constitution soon directed its attention to Reverend Long’s conspicuous lifestyle, including his $3.1 million salary and his $1.4 million home, all of which were provided by his Church, along with a Bentley automobile worth $350,000.35 The lavish lifestyle is not, by itself, sufficient to conclude that charitable insider trading is occurring. The report continues though by describing a web of 20 for-profit and non-profit entities owned by Reverend Long and through which, the report suggests, Reverend Long is compensated for transactions with the 25,000 member New Birth Missionary Baptist Church.36 In effect, Rev. Long is portrayed as monopolizing the very considerable consumer power for his personal profit. Although the article quotes several sources for the proposition that Reverend Long obviously overcharged his Church and the other nonprofit organizations for which he serves as an insider, it sets forth no evidence in support of that factual conclusion. The whiff of scandal remains, otherwise the Atlanta Journal-Constitution would not have been interested in the story. Assume that the innuendo in the report is true to the extent that Reverend Long causes the New Birth Missionary Baptist Church – that is, the congregation – and his other unspecified nonprofit organizations to transact business on fair terms, but solely with his for-profit entities. If indeed that were true donors would be legitimately concerned that their patronage is being used for an insider’s personal profit rather than on behalf of the beneficiaries on whose behalf the organizations purportedly exist. It would be nearly impossible, though, to condemn or even be concerned with the transactions under present

36 Id.
law or the Service’s recent proposal because all the transactions involved reasonable transfers and presumably there is a sufficient amount of real charity occurring as to immunize the transactions from legal challenge.\footnote{As described in greater detail below, the proposed implementation of joint venture private inurement allows an insider to escape liability if the organization nevertheless primarily engages in legitimate charitable activity even as the insider skims off the top.}

In short, there is yet no apparent “result” with which the law is concerned and since the law is unconcerned with process, even process that inevitably causes harm, there is simply no legal violation. The theory of joint venture private inurement, as offered in my previous article and as described in greater detail in the following sections, would provoke the suspicion and justifiable hostility that would make an actual result unnecessary and thereby better protect the market’s reputation. In the Bishop Long case, the damage to the nonprofit market reputation is complete. Present law and the Service’s proposed amendment provide no recourse either to prevent or respond.
by Wendy C. Gerzog

Last spring, the IRS issued Rev. Proc. 2005-24, applicable to the qualification of inter vivos charitable remainder annuity trusts (CRATs) and charitable remainder unitrusts (CRUTs) where the surviving spouse can satisfy her statutory elective share from the assets of a charitable remainder trust (CRT). Section 664(d)(1)(B) defines a CRAT as allowing no payments other than qualifying annuity and certain other payments to a person other than a qualified charity. Section 664(d)(2)(B) defines a CRUT similarly. The Rev. Proc. was intended to provide a safe harbor rule under which the IRS would ignore the elective right in its determination of a trust’s qualification and continuation as a CRT. Extremely unpopular and creating a great deal of negative criticism, particularly from charities, the IRS recently extended its grandfather date.

---

2 Under section 664(d)(1)(A), an annuity consists of “a sum certain (which is not less than 5 percent nor more than 50 percent of the initial net fair market value of all property placed in trust) [which] is to be paid, not less often than annually, to one or more persons (at least one of which is not an organization described in section 170(c) and, in the case of individuals, only to an individual who is living at the time of the creation of the trust) for a term of years (not in excess of 20 years) or for the life or lives of such individual or individuals.”
3 Under section 664(d)(1)(C), after the termination of the annuity payments, “the remainder interest in the trust is to be transferred to, or for the use of, an organization described in section 170(c) or is to be retained by the trust for such a use or, to the extent the remainder interest is in qualified employer securities (as defined in subsection (g)(4)), all or part of such securities are to be transferred to an employee stock ownership plan (as defined in section 4975(e)(7)) in a qualified gratuitous transfer (as defined by subsection (g)).”
4 Instead of a retained annuity of “a sum certain,” a CRUT is defined as “a fixed percentage... of the net fair market value of its assets, valued annually.” I.R.C. §664(d)(2)(A).
5 See, e.g. Comments on Revenue Procedure 2005-24 of the American Bar Association Section of Real Property, Probate and Trust Law, Kevin L. Shepherd, Chair, 2006 TNT 12-19 [hereinafter RPPT Comments] (“Understandably, the Service wants to protect assets placed in CRTs from the claims of spouses. We believe, however, that the Revenue Procedure may chill the climate for the creation of CRTs without providing any significant additional protection for charitable assets.” Id.); Comments of Judith W. McCue, President, American College of Trust and Estate Counsel (ACTEC), Tax Analysts Doc. 2006-253(letter to the Hon. Mark W. Everson, Commissioner, IRS, dated November 22,2005 and received December 15, 2005)[hereinafter, ACTEC letter] (“The additional expense and inconvenience of compliance with Rev. Proc. 2005-24 will have the net result of discouraging donors from establishing CRTs or cause them to contribute less.” Id. at 6.).
6 See, e.g. e-mail and attached letter from Frank Minton, The American Council on Gift Annuities, a 501(c)(3) organization sponsored by about 1,250 educational, religious, welfare, health care, and
until further notice. For grandfathered trusts, the IRS has announced that it “will disregard the right of election, even without a waiver, but only if S [the surviving spouse] does not exercise the right of election.”


States with UPC elective share provisions may allow the surviving spouse to elect a percentage of the “augmented estate” that includes property transferred by the deceased to others, such as CRT assets. According to Rev. Proc. 2005-24, because of the possibility that she may choose to apply her right against CRT property, the election right itself for non-grandfathered trusts, or the exercise of that right for grandfathered trusts, results in the failure of a CRT under section 664(d). In order to correct this defect and in order for the trust to obtain and retain its exempt status as a CRAT or CRUT, the IRS proposed a safe harbor by which the grantor’s spouse must irrevocably waive her right to apply her spousal share against CRT assets. This waiver has to be filed within six months after the Form 5227 due date for when the later of the following events occur: the trust’s creation, the grantor’s marriage, the date the grantor is domiciled or resides in a jurisdiction with the applicable elective share provision, or

---

8 The June 28, 2005 grandfather date was indefinitely extended. *Id.*
10 The percentage varies depending on the length of the couple’s marriage from the “supplemental amount only” to 50 percent of the augmented estate for those couples married for fifteen years or more. U.P.C. §2-202 (Rev. 1990).
11 U.P.C. §2-203 defines the composition of the augmented estate to include “the decedent’s nonprobate transfers to others.” The augmented estate includes trusts created by the decedent with retained income, annuity, or unitrust interests, for life, for a term of years, or relinquished within two years of his death. See U.P.C. 2-205, Comment, Examples 9, 10, 12,16.
12 In community property states, in states which have not adopted the concept of the UPC augmented estate or in states whose augmented estate does not include CRTs, naturally, no waiver is required under the Rev. Proc. Rev Proc. 2005-24, at *5.
13 This due date for filing the “Split-Interest Trust Information Return” does not include extensions granted. *Id.* at *7.
the date the grantor’s state effectuates such a law. A copy of the spouse’s signed waiver must be provided to and preserved by the CRT’s trustee.

**Qualification as a CRT**

The words of the statute are clear: the definitions of CRAT and CRUT prohibit transfers to any person other than those with the applicable retained annuity or unitrust interest or the permissible remainder beneficiaries. It is apparent by the language of section 664(d) that Congress was directing how a CRT instrument must be drafted. The language, especially the parenthetical expressions, of the CRAT definitional statute lays out very specific requirements for the retained non-charitable annuity interest in the trust. The annuity must be:

- a sum certain (which is not less than 5 percent nor more than 50 percent of the initial net fair market value of all property placed in trust) [which] is to be paid, not less often than annually, to one or more persons (at least one of which is not an organization described in section 170(c) and, in the case of individuals, only to an individual who is living at the time of the creation of the trust) for a term of years (not in excess of 20 years) or for the life or lives of such individual or individuals.

Similar language with mostly the same specific requirements applies also to the noncharitable unitrust interest of a CRUT instrument.

Moreover, the statutory exception within this definitional section states:

Notwithstanding the provisions of paragraphs (2)(A) and (B), the trust instrument may provide that the trustee shall pay the income beneficiary for any year—

(A) the amount of the trust income, if such amount is less than the amount required to be distributed under paragraph (2)(A), and

(B) any amount of the trust income which is in excess of the amount required to be distributed under paragraph (2)(A), to the extent that (by reason of subparagraph (a)) the aggregate of the amounts paid in prior years was less than the aggregate of such required amounts. [emphasis added]

---

14 *Id.* at *7-8.
15 *Id.* at *8.
18 I.R.C. §664(d)(3).
Certainly, if the governing instrument itself provided for the surviving spouse to have a contingent interest in the CRT, the trust would not qualify as a CRT.

However, some questions remain: 1) whether the statutory prohibition against additional non-charitable remainder beneficiaries was intended to apply to fluctuations of local law; 2) whether the purposes behind the statutory restrictions are carried out through Rev. Proc. 2005-24; 3) whether the numerous triggering events are administratively unworkable and whether a recapture provision at decedent’s death (one which would account for not only the charitable deduction originally taken by the grantor but also the increased value of the corpus for undistributed income erroneously allowed to be tax-exempt) would be more practicable; and 4) whether the revenue procedure manifests a special aversion to the UPC augmented estate elective share and, together with some commentaries, has sexist results.

A. Connection between Local Law and Qualification for Tax Benefits

There are numerous instances where local law property characterizations affect qualification for tax benefits, particularly in the transfer tax area. For example, in *Jackson*, the Supreme Court disallowed a marital deduction for a widow’s allowance that, under state law, was defeasible on her death or remarriage. Because it would “terminate or fail” on the occurrence of either of these events and would thereby pass to a trust for his daughter, as determined from the date of the decedent’s death, the widow’s allowance under the California statute did not qualify for the marital deduction.

---

20 *Jackson v. United States*, 317 F.2d 821, 826 (9th Cir. 1963) (“In this case the interest which passed from the decedent to his widow was in the corpus of the decedent's estate in which another interest had passed to decedent's daughter. If for any reason the widow's right to the marital allowance failed, it would not have gone in all events to her estate but it would have served to augment the residuary estate in which the widow had only a limited and terminable interest. In such event a part of such property could be possessed and enjoyed by the daughter.”)
21 The court held that qualification for the marital deduction had to be determined at decedent’s date of death. *Jackson*, at 508.
Other state widow’s allowance statutes that create vested rights and are not defeasible qualify for the marital deduction.\textsuperscript{22} Likewise, where the widow is decedent’s sole beneficiary, the interest qualifies for the marital deduction because on the event of its failure, it would nevertheless devolve to her.\textsuperscript{23}

There is some instruction from \textit{Jackson} to the situation in the Rev. Proc. Here, we have a local law dealing with marital rights that potentially will deprive the charity of its remainder interest. If the spousal election is made, the grantor will have received an unwarranted income and gift tax charitable deduction and undistributed trust income and gains will erroneously have been allowed to go untaxed. Here, however, unlike the date of death analysis of the widow’s allowance as a terminable interest in \textit{Jackson}, we don’t have just one date to determine the qualification of the CRT. The Rev. Proc. requires a waiver six months after the later of any of the following four triggering events: the trust’s creation, the grantor’s marriage, the date the grantor is domiciled or resides in a jurisdiction with the applicable elective share provisions, or the date the grantor’s state effectuates such a law. Since it is likely the grantor would have to file waivers at each of these events, it is administratively unworkable.

\textbf{B. Policy of CRT Rule and Unjust Enrichment}

The policy behind the split-interest rules for gifts to charities was that Congress wanted to ensure that “there is a reasonable correlation between the amount of the deduction and the

\textsuperscript{22} See, e.g., \textit{Estate of Watson v. Commissioner}, 94 T.C. 262 (1990)(“Under Mississippi law, a widow has an absolute right to the widow's allowance provided for in Miss. Code Ann. Section 91-7-135 (1972). [citations omitted] The possibility that, as a practical matter, a widow could lose that right if the appraisers fail to set aside the widow's allowance and the widow fails to claim the allowance before the decedent's estate closes does not change the fact that the widow had an absolute right under Miss. Code Ann. Section 91-7-135 (1972) at the date of the decedent's death.” \textit{Id.} at 280.).

\textsuperscript{23} See Reg. \textsection 20.2056(b)-1(g) \textit{Example 8} (“Under such circumstances, the allowance constitutes a deductible interest since any part of the allowance not receivable by the surviving spouse during her lifetime will pass to her estate under the terms of decedent’s will.”).
benefits that the charity will ultimately receive.” Pre-1969 law allowed deductions of remainder interests in trust that did not reflect the charity’s actual benefit because trust assets could be invested to augment the income interest so that there was “little relation between the interest assumptions used in calculating present values and the amount received by the charity. For example, the trust corpus can be invested in high-income, high-risk assets. This enhances the value of the income interest but decreases the value of the charity's remainder interest.” Other instances of overvaluation of the charitable interest included “where the charity has only a contingent remainder interest in the trust (for example, a $5,000 annuity to A for life, remainder to his children, or to a charity if A has no children)” or “where a charity has a remainder interest and the trust permits invasion of the charitable share for the benefit of a noncharitable intervening interest which is incapable of reasonably certain actuarial valuation (for example, a $5,000 annuity to A for life, remainder to a charity, but the trust provides that the trustee may pay A amounts in excess of $5,000 in order to maintain his standard of living).” Congress

---

24 S. Rep. 91-552, Pub.L. 91-172, Tax Reform Act of 1969, at 2114 (Nov. 21, 1969). That this section is concerned with accurate valuation of the split interests in the trust has been underscored by its imitation in the enactment of Chapter 14, “the new valuation rules,” enacted in 1990 and applied to such trusts as GRATs and GRUTs, non-charitable split-interest gifts. The Revenue Reconciliation Act of 1990 added sections 2701-2704 as part of the amendment to Subtitle B, which added Chapter 14. See Revenue Reconciliation Act of 1990, Pub. L. No.101-508, 104 Stat. 1388-1491 (1990) (codified as amended at I.R.C. §§2701-2704). Section 2702 is effective for transfers after October 10, 1990, except as provided in section 11602(e)(1)(B) of the amendments. See id., 104 Stat. 1388-1501. The problem of terminable interests, estate freezes, and GRITs is the undervaluation of property transferred to third parties. According to the legislative history of section 2702, “the committee [was] concerned about the undervaluation of gifts valued pursuant to Treasury tables. Based on average rates of return and life expectancy, those tables are seldom accurate in a particular case, and therefore, may be the subject of adverse selection. Because the taxpayer decides what property to give, when to give it, and often controls the return on the property, use of Treasury tables undervalues the transferred interests in the aggregate, more often than not. Therefore, the committee determines that the valuation problems inherent in trusts and term interests in property are best addressed by valuing retained interests at zero unless they take an easily valued form—as an annuity or unitrust interest. By doing so, the bill draws upon ... rules valuing split interests in property for purposes of the charitable deduction.” 136 CONG. REC. 30538-39 (1990).
25 Id. at 2116.
26 Id. at 2117.
27 Id.
restricted the form of a charitable remainder interest in trust to either a CRAT or a CRUT.\textsuperscript{28}

Essentially, Congress was most concerned with the grantor’s (or trustee’s) use of a trust to manipulate and reduce or eliminate the amount that would actually be paid to a charity while at the same time using the calculations of the actuarial tables to compute the grantor’s charitable deduction. The provisions are meant to define the retained income interest more rigidly to make the computed value of the charitable gift reflect more of its actual value.\textsuperscript{29}

A grantor receives income\textsuperscript{30} and gift\textsuperscript{31} tax deductions when he creates a CRT and transfers property to the CRT. Taking those deductions for amounts that do not actually go to a charity clearly contravenes the policy behind the CRT statutes. Moreover, the trust has been allowed to grow without being subject to income taxes for undistributed income and gains. That too flouts the purpose of section 664 and the charitable split-interest rules.

C. Issues of Timing and Marital Status

With respect to CRTs, qualification as an inter vivos CRAT or CRUT is necessary in order to obtain income and gift tax deductions as well as to remain tax-exempt during its term. At the time the grantor transfers assets to the CRT, the trust instrument can be

\textsuperscript{28} That is, Congress decided that “a deduction would not be allowed for a gift of a remainder interest in trust to charity unless the gift took a specified form: namely, an annuity trust (under which the income beneficiary is to receive a stated dollar amount annually) or a unitrust (under which the income beneficiary is to receive an annual payment based on a fixed percentage of the trust's assets).” \textit{Id. at} 2116.

\textsuperscript{29} That is not to say that with the CRTs, there is no abuse. Clearly, CRTs are part of the estate planner’s tools to transfer property to one’s family and friends without incurring the actual matching gift or estate taxes for these transfers. Thus, a CRAT is advised for in poor economic times where the actual investment return is less than the amount of the annuity transferred. CRUTs, NIOCRUTs, NIMCRUTs, and flip CRUTs can be established to provide transfer tax savings. However, the 1969 legislation imposed some boundaries on the extent of that manipulation and those plans are not without their own risks or additional costs. The choice of CRT, depending on the economy and the trust’s actual investments, may not produce the expected benefits. \textit{See} Daniel Halperin, \textit{A Charitable Contribution of Appreciated Property and the Realization of Built-In Gains}, 56 Tax L.Rev. 1, 13 (2002)(“Given the $4 tax savings from the transfer of a $10 interest to charity and the value of the noncharitable interest in the trust ($90), Charlotte effectively may obtain tax-free diversification while giving up only 6% of the value of the property to charity. This behavior might make sense even for a donor who has no charitable impulse. Since there is very little benefit to charity, the inequity of allowing diversification without gain recognition is particularly strong in the case of a partial transfer and seems unwarranted even if the treatment of outright gifts is unchanged.” \textit{Id.})


\textsuperscript{31} I.R.C. §2522(c)(2)(A).
examined to determine whether the trust language comports with the requirements of section 664. But, what about local law which, as the Rev. Proc. acknowledges, may change over time? Also, what about the grantor’s marital status that, again, may shift—perhaps numerous times—over his lifetime? Are there any other tax benefit statutes that require waivers at the happening of several future events?

There is something ridiculous about the scenarios engendered by this Rev. Proc. The grantor will need a waiver from his spouse when he moves thirty years after the CRT’s creation into a state that has a UPC augmented share statute. The grantor who got a waiver from wife #1 will need successive waivers from wives #2 through 4, but only if he remains in or moves to a state that currently has an applicable UPC augmented share statute. Even if the grantor remains in the same state and stays married to the same wife, he, or more likely the trustee, needs to make the connection between changes in local law and the CRT he may have created many years ago.

The Rev. Proc. makes what is an inter vivos CRT dependent for qualification on events that will occur, if they occur at all, shortly after the grantor’s death. It is true that if the surviving spouse makes an election that applies to the charitable interest in the trust, the charity will not benefit. But, perhaps it should only be at that time, after the decedent’s death when the spousal election is applied to the CRT, that a recapture provision should be given effect.

32 Besides at the trust’s creation, the Rev. Proc. requires that the spouse file a waiver on the following dates: the grantor’s marriage, the date the grantor is domiciled or resides in a jurisdiction with the applicable elective share provisions, or the date the grantor’s state effectuates such law. Rev. Proc. 2005-24, at *7-8.
33 Relatedly, must he also inform the trustee of each marriage, divorce, or change of residence? See ACTEC letter, at 5.
The Rev. Proc. is unclear what must be done if the waiver is not submitted. Some commentators have suggested a recapture provision, including legislation that would require the creator of a CRT the assets of which are diverted from charity to include in his or her income tax return for the year in which the diversion occurs, or in his or her final return, if diversion occurs after death, an amount equal to the amounts deducted under Section 170 for the contribution to the CRT and an amount equal to the amount that would have been included in his or her gross income during the term of the CRT if the CRT had been treated as owned by him or her within the meaning of Section 671 of the grantor trust rules.

Certainly, a recapture provision must recover all of the benefits associated with the trust’s having been erroneously classified as a CRT: the income and gift tax charitable deductions and the income and gains that were untaxed because the CRT was tax-exempt.

To illustrate: Decedent D died in Colorado, a state that has adopted the UPC elective share and augmented estate, and has an abatement statute that applies in cases of a spousal election. During life, D established a CRAT, retained an annuity for his life, and received $1 million in tax benefits. The CRAT is worth $5 million at D’s death.

---

34 See RPPT Comments, supra note 5 (Pt.III) (“For example, if the statute of limitations for the tax return on which the CRT’s creator claimed a charitable deduction for his or her gift to the CRT is open, will the creator be required to file an amended return to reflect the loss of the deduction? If the statute of limitations has run on the creator’s return on which he or she claimed a charitable deduction and on the CRT’s return that reported a capital gain on the sale of the originally, contributed assets to the CRT, what will be the Service’s remedy?” Id.); “There are potentially draconian consequences (such as loss of the donor’s income, gift, and estate tax charitable deductions and loss of the CRT’s tax-exempt status) if the donor fails to comply with the safe harbor spousal waiver, whether such failure is knowing or unknowing and whether presently or in the future. *It is impossible (notwithstanding the ‘safe harbor’ label) to know or ensure that a trust will be a qualified CRT as of the date of its inception and on an ongoing basis, regardless of changes in circumstances. This means that a donor cannot rely on the income, gift, and estate tax status of a CRT in other tax and estate planning matters.” ACTEC letter, supra note 5, at 4.
35 RPPT Comments, supra note 5 (Pt.IV).
36 Some have, however, questioned whether “the problem identified in Rev. Proc. 2005-24 [is] worth the potential damage to the statute of limitations?” ACTEC letter, at 7. If the spousal election is made with respect to CRT assets, such diversion from the charity should require a remedy, such as the recapture of the grantor’s benefits.
and D left S, his son, $5 million in his will. W, who had been married to D for 15 years before his death, elects her 50% share of D’s augmented estate, which includes the CRAT. Under a recapture solution, the government receives $1 million from D’s estate and W receives one-half of D’s net estate and, if the charitable transfer is contingent on the deduction, S receives the other half.


1. Policy of UPC Augmented Estate

The UPC elective share revisions of the 1990’s were intended “to bring elective-share law into line with the contemporary view of marriage as an economic unit,” reflective of the treatment of the couple’s property division in a divorce in both community property and common law property states. One of the purposes of the redesigned elective share is “to reward the surviving spouse who sacrificed his or her financial-earning opportunities in order to contribute so-called domestic services to the marital enterprise and to deny an additional windfall to the surviving spouse in whose name the fruits of a long-term marriage were mostly titled.”

According to the UPC comment, “The pre-1990 Code made great strides toward preventing ‘fraud on the spouse’s share.’” That fraud occurs where decedent makes inter vivos transfers for the purpose of diluting or eliminating the spouse’s elective share.

---
39 The surviving spouse’s share will most likely qualify for the estate tax marital deduction under I.R.C. §2056. Property that she receives from her deceased spouse by the elective share qualifies as “passing” to her under I.R.C. §2056(c)(3)(“For purposes of this section, an interest in property shall be considered as passing from the decedent to any person if and only if—(3) such interest is the dower for curtesy interest (or statutory interest in lieu thereof) of such person as surviving spouse of the decedent; . . .”).
40 But see Proctor
41 If the charity does not have to return any excess, S and the charity would each receive one-fourth.
42 “The elective share of the surviving spouse was fundamentally revised in 1990 and was reorganized and clarified in 1993.” UPC, Part 2: Elective Share of Surviving Spouse, General Comment.
43 Id.
44 Id.
45 Id. (“The Redesigned Elective Share”).
46 Id. (“Decedent’s Nonprobate Transfers to Others”).
right. As a solution, the pre-1990 Code adopted the concept of the augmented estate.

With the revisions to this part of the UPC, “the augmented-estate concept has been strengthened.” Likewise, courts have expressed the rationale for the augmented estate concept of the UPC was intended to protect the surviving spouse from her propertied spouse’s inter vivos transfers to third parties and to fortify the concept of the elective share.

With respect to inter vivos CRTs, the augmented estate is limited to decedent’s transfers during marriage: 1) wherein he retained an income right from the property that terminated at his death or afterwards; or 2) wherein he created a power over income or principal to benefit himself, his creditors, his estate, or his estate’s creditors. Further, the augmented estate applies to CRTs where, during his marriage and within two years of his death, the decedent transferred property to anyone other than the his surviving spouse (to the extent that the total transfers to that person in either of the two years was more than $10,000).


---

47 Id.
48 See, e.g. In re Carmen’s Estate, 213 Neb. 98, 100, 327 N.W.2d 611, 613 (1982) (“The combined effect of the statutory elective share and augmented estate concepts is intended to protect the surviving spouse of a decedent against donative inter vivos transfers by devices which would deprive the survivor of a ‘fair share’ of the decedent's estate and at the same time prevent the surviving spouse from receiving more than such share by allowing the acceptance of certain transfers and insurance proceeds and also yet elect against the will.” Id.); Matter of Estate of Smith, 718 P.2d 1069, 1071 (Colo. App. 1986)(“We perceive that one of the purposes of the augmented estate provisions is to allow a surviving spouse, deprived of a share in the decedent spouse's estate by inter vivos transfers to third parties or other means, to claim an elective share of property deemed includible in the augmented estate.” Id.).
51 That is, the augmented share applies either to a CRAT created within two years of decedent’s death or to the creation or transfers to a CRUT within those two years. Additional transfers to CRATs are prohibited after the initial transfer to the trust; however, additional contributions may be made to a CRUT. Reg. §§1.664-2(b) (“A trust is not a charitable remainder annuity trust unless its governing instrument provides that no additional contributions may be made to the charitable remainder annuity trust after the initial contribution.”); 1.664-3(b)(1) and (2) (additional contributions may be made subject to additional valuation and computation requirements).
52 UPC §2-205(3)(iii) (2005).
Some of the comments and criticisms of the Rev. Proc. have pointed out its under-inclusiveness. “For example, a CRT created by an individual who is a domiciliary of a civil law country could become subject to forced heirship claims.” 53 Likewise, in Louisiana, CRTs may be subject to decedent’s children’s claims. 54 Outright charitable gifts may also be subject to spousal election rights, 55 especially if decedent made gifts within two years of his death, 56 and CRT assets may be subject to creditors’ claims under bankruptcy law. 57

Likewise, some have questioned the effect of the Rev. Proc. on “other split-interest charitable gifts (e.g., pooled income funds and remainder interests in personal residences or farms) that may also be subject to spousal election rights and thus ultimately involve similar requirements and problems.” 58 Further, they have speculated on the consequences of the Rev. Proc. on “all gifts, including outright gifts and non-charitable gifts, that may also be subject to spousal election rights and thus be incapable of being completed gifts (at least without a spousal waiver, or even with a waiver, if new spousal election rights could emerge upon remarriage or change of domicile).” 59

Therefore, some of the comments suggest that the IRS ignore this contingency. 60 They point to the fact that few even make the election, 61 exercising it against the assets in

53 RPPT Comments, supra note 5, (Pt. I).
54 Id., referring to both New York and Virginia and citing New York law (New York Estates, Powers & Trusts Law §5-1.1-A(b)(1)(B)).
57 ACTEC letter, supra note 5, at 6.
58 Id. at 6-7.
59 “Thus, ACTEC believes that the Service is justified in concluding that the likelihood of any particular spouse actually taking assets of a CRT pursuant to an elective share right is so remote as to be negligible and may be ignored as a contingency which would cause the CRT to fail to meet the definition of and function exclusively as a CRT from inception.” Id. at 7. ACTEC recommended, therefore, that Reg. §1.664-1(a)(4) be amended to conform with its belief and suggested additional regulatory language. Id. at 7-8.
a CRT is probably extremely rare.\footnote{Logically, that means, either men provide well for their wives or the widows do not get adequate advice about the advantages of, or are otherwise inhibited from, making the election.} “Anecdotal evidence indicates that the possibility of the exercise of a spousal right of election against a CRT is remote. The authors of these comments have been involved, collectively, in the creation of thousands of CRTs and not one of them has ever been aware of a spousal exercise of a right of election against a CRT.”\footnote{RPPT Comments, \textit{supra} note 5. Thus, the comment considered the spousal election to be “a contingency so remote as to be negligible.”} So, why single out the contingency of the UPC augmented estate elective share?

3. Sexism

Since women are more often the “surviving spouse,” most augmented share statutes benefit widows. Thus, the sexist consequence of targeting the spousal election right is disturbing. Moreover, some of the reactions to the Rev. Proc. ironically reflect the need for the spousal election right against the augmented estate.

Because a waiver may be unenforceable in many states or may require complete disclosure and separate spousal representation, a grantor may decide not to create a CRT rather than go to “the expense and emotional difficulties of negotiating a post-nuptial agreement with his or her spouse.”\footnote{Id. (Pt.IIA). See ACTEC letter, \textit{supra} note 5, at 3.} Indeed, states that have adopted the UPC augmented estate concept are those particularly likely to inquire about the circumstances of the surviving spouse’s waiver and to insist that the waiver was a knowing one.\footnote{\textit{Id. See, e.g., In re Estate of Smith, 674 P.2d 972 (Colo. App. 1983)(“The right to statutory allowances under §§15-11-201, 15-11-402, and 15-11-403, C.R.S. 1973 (1982 Cum. Supp.) is strongly favored in Colorado, and that right may not be held to have been waived or relinquished except in cases where the specific intention to waive or relinquish is established by the evidence. See McLaughlin v. Craig, 117 Colo. 67, 184 P.2d 130 (1947), interpreting a similar provision in the predecessor statute to §15-11-201, C.R.S. 1973, and requiring waiver to be established beyond any reasonable doubt.” \textit{Smith}, at 973.)} Such a comment, however, suggests the chill of full disclosure is rooted in precisely the reasons that a surviving spouse should \textit{not} want to waive her election right.
Likewise, some of the recommendations place the burden on the surviving spouse instead of on the grantor. “If it is decided not to withdraw the Revenue Procedure, we recommend the addition of an alternate safe harbor that relies on penalties that are likely to be imposed under Section 4941 on any spouse who exercises his or her right of election.”66 That is, “A spouse faced with the prospect of a 200% penalty on the amount paid to him or her by a CRT is not likely to exercise his or her right of election against the CRT.”67

First, the Rev. Proc. requires a spousal waiver at a time when she may not be able to weigh consequences that won’t be clear until her spouse’s death, perhaps many years after the establishment of the CRT. Then, there is the comment that if she ever tries to do anything so heinous as to consider applying her right of election against a CRT, she’ll be subject to large penalties. “As a result, if he or she is informed of the existence of this penalty, the possibility that he or she would exercise his or her right of election against the CRT is insignificant.”68

Rather than suggest that the husband should adequately provide for his widow so that she would not want to take her elective share and so that his estate would not be liable for all of the recaptured tax benefits that he derived from the CRT, the comments emphasize the penalties that will, and should, be asserted against the widow should she attempt to make that election.

---

66 Id. (Pt.IIB). The section 4941 penalty is imposed against self-dealing for transferring private foundation assets to a disqualified person. Id.
67 Id.
68 Id. “In our experience, many spouses are often wary when asked to sign any waiver of rights, even if there are no rights to waive or if the exercise of such rights is effectively prohibited by federal tax law (see discussion of Code §4941 below.” ACTEC letter, at 4.
The penalties include those under Code sections 507, 4941, and 4945. These penalties apply because a CRT is treated as if it were a private foundation. Penalties against self-dealing pertain to acts of the relatives of the creator; however, they apply because there is an assumption that their interests are the same. In the case of the spousal election, their interests are clearly averse to one another. So, while the penalty statute literally applies to the widow who makes the UPC election against the decedent’s augmented estate, the logic of the identity of interests is absent.

The foundation penalty statutes do not apply to “any amounts in trust other than amounts for which a deduction was allowed . . . if such other amounts are segregated from amounts for which no deduction was allowable. . . .” Indeed, some of the recommendations specifically reference, and are reactions to, the government’s fear of a “whipsaw” effect if the surviving spouse exercises her election, with a claim that the CRT had never qualified as such and, thus, no deduction was allowable for the transfers to the CRT.

69 See ACTEC letter, at 8-15.
70 I.R.C. §4947(a)(2).
71 I.R.C. §4941(d).
72 I.R.C. §4940(d)(3)(B)(iii) (“(B) Disqualified individual.--The term "disqualified individual" means, with respect to any private foundation, an individual who is--
(i) a substantial contributor to the foundation,
(ii) an owner of more than 20 percent of--
(I) the total combined voting power of a corporation,
(II) the profits interest of a partnership, or
(III) the beneficial interest of a trust or unincorporated enterprise, which is a substantial contributor to the foundation, or
(iii) a member of the family of any individual described in clause (i) or (ii).”).
74 By amending the regulations, treasury will prevent the “whipsaw” effect by the surviving spouse to avoid penalties on the basis that the CRT never qualified as such. ACTEC letter, supra note 5, at 8, n.7. See e-mail from Robert Rosepink to Catherine Hughes (July 12, 2005), Doc. 2005-16114, Taxanalysts Document Service (“If I understood you correctly, the Service is concerned about being whipsawed by a taxpayer who argues that if the CRT isn’t a qualified CRT under section 664 since inception, then section 507 can’t be invoked.”).
If a recapture provision applies to the grantor, it will indicate that no deduction was allowable because of the contingency of the spousal election applicable to the UPC augmented share. A recapture provision is, therefore, preferable to threats of penalties as it recovers the tax benefits from the CRT’s creator (or his estate) rather than singling out his widow for punishment75 because he did not adequately provide for her in his will.

Some question whether “the problem identified in Rev.Proc. 2005-24 [is] worth the potential damage to the statute of limitations.”76 However, there are many recapture sections in tax, including recapture sections applicable to estate tax.77 Also, since the CRT doesn’t distribute its assets to the charity until the termination of the non-charitable interests, generally at the creator’s death,78 the charity will not have been able to exhaust trust assets before recapture.

Conclusion

The IRS should “disregard the right of election, even without a waiver, but only if S [the surviving spouse] does not exercise the right of election.” In such an event, the IRS should impose a recapture requirement that would encompass recovering the benefits the grantor erroneously received because of the trust’s qualification as a CRT. This course of action preserves the charitable deduction from being diverted to the surviving spouse in contravention of section 664 and, at the same time, it is a more workable solution that will avoid what could well be some absurd consequences of the 2005 Rev. Proc.

Likewise, the IRS can temporarily ignore the contingencies of forced heirship on CRTs

75 Admittedly, the pot from which she will be able to apply her elective share will be a smaller one after the taxes are recaptured; however, it will merely place her in the position she would have been in had the charitable deductions been earlier denied to the CRT’s creator. While her share may thus be reduced, so will the third party beneficiaries of the decedent’s transfers.
76 ACTEC letter, at 7.
77 See, e.g. I.R.C. §2032A(c)(“Tax treatment of dispositions and failures to use for qualified use”).
78 For the UPC augmented estate elective share to apply, the creator would have had to retain an income interest until his death or within two years of that time. See UPC §§2-205(2)(i); 2-205(3)(iii) (2005).
and other charitable gifts that are subject to the UPC augmented estate elective share statutes until the surviving spouse actually exercises her rights against property.

This recommendation also has the advantages of respecting one of the main purposes of the UPC’s augmented share and of not pressuring mostly women to relinquish rights when they cannot appreciate the consequences of the waiver. In reaching its decision in *Jackson*, the Supreme Court concluded, “The achievement of the purposes of the marital deduction is dependent to a great degree upon the careful drafting of wills; we have no fear that our decision today will prevent either the full utilization of the marital deduction or the proper support of widows during the pendency of an estate proceeding.” Although the Rev. Proc. may not affect many widows and the Rev. Proc. might benefit some widows from the full disclosure required for an effective waiver, eliminating the waiver requirement will complement the UPC’s augmented share goal by not prematurely restricting that elective share right.

Moreover, the focus with a recapture section should be on the creator of the CRT who did not sufficiently provide for his surviving spouse. Since CRTs are often prepared in connection with the creator’s estate plan, it would be best at that time to explain to the propertied spouse and creator of the CRT that all of the tax benefits will be recaptured if he does not adequately provide for his spouse in his will (or otherwise). That advice, if taken, should obviate her desire to elect her spousal share. At any rate, certainly, that solution would be preferable to applying extreme pressure (in terms of penalties) on the widow to inhibit her from making her spousal election.

---

*Jackson*, at 511.
Untapped Wealth Maximization: The Under Utilization of the 501(c)(3) Charitable Entity to Level the Playing Field in Urban Communities of Color

By Alice M. Thomas

I. Introduction

By addressing myself to the role of nonprofit organizations in building wealth for low income people in urban communities of color, I am essentially exploring the role that government should play through its tax policies in enabling families to build wealth. By building wealth, families can increase their mobility, stability and well-being. The government does have such a role, and the 501(c)(3) charitable entity is a viable vehicle to further government public policy in this arena. The strategies suggested in this article are at the individual, family, and community level. At the community level, collective will is expressed through organizations working in the community for change. Various kinds of nonprofits have worked for change in communities of color for more than 100 years. Following the freeing of hundreds of thousands of Africans from bondage, benevolent societies, churches and other community organizations filled in the gap between governmental provision of resources and private action. The more I explore the depths of the problem, the more painfully obvious it becomes that no one intervention will suffice. A multilayered strategy is needed to address this catastrophic problem. It is such a big problem because wealth is a major indicator for a plethora of ills that do and can befall an individual, family or community in the absence of wealth. The 501(c)(3) public charity is uniquely situated to assist in leveling the playing field for persons traditionally overlooked in federal tax policy.

In the last ten years, there has been an increasing awareness of the role of tax policy in asset accumulation strategies for low-income persons. While there are more whites who are low income in absolute numbers than persons of color, there is a higher percentage by group of persons of color that are low income. As a consequence, any strategy that targets the low income also stands to largely benefit communities of color. For ease of argument, my primary focus is on the African American community because most if not all of the strategies would apply equally well in Latino communities.

The role I want nonprofits to play in leveling the playing field is triplefold – garnering or creating funds for wealth building initiatives, providing financial education, and transferring

1Associate Professor of Law at the University of the District of Columbia David A. Clarke School of Law.

2List Interventions
The target population is about 50 percent of all African Americans. About 40 percent of African Americans live at or below the poverty line, and African Americans as an aggregate possess about seven cents on the dollar of the wealth possessed by white Americans, and Latinos possess about eleven cents on the dollar. (Need source)


R. Kuttner, *Sharing America’s Wealth: the Policies and Politics of Building a Larger Middle Class* (http://www.prospect.org/kuttner-r-sr.html). L. Schweninger, Black Property Owners in the South 1790-1915 (This book documents the patterns of black property ownership among slaves and free men and women over a hundred year period. The author documents that, despite the legal and societal impediments to black property ownership, blacks’ attitudes toward individual property ownership were intense, and led to greater numbers than anticipated owning land and business enterprises. *Id.* at 236-36. The author further states that, “Despite slavery, racism, war, emigration, colonization, lynchings, and brutal murders, most blacks continued to view the South as their home and clung to the values they had grown to accept: that acquiring land and property would somehow free them from the burdens of the past. Their tragedy, and the region’s tragedy, was that it never would.” *Id.* at p. 237.
For example, the burning of _______, a black town in Tulsa, Oklahoma where blacks were way ahead of the curve in wealth accumulation. (Look up this reference), and (reference the Florida city, here.) These towns were self-sufficient enclaves of wealth amassing African American families at the turn of the century. The playing field at the economic level was leveling off. After the destruction by mayhem, plundering and fire, these communities of African Americans never rebounded, and the wealth advantages were lost forever to generations of families. Today, there are few enclaves of concentrated populations of African Americans who are financially secure. Prince George’s County in Maryland, a majority African American county, is touted as having the largest concentration of high income African Americans in the country. I do not have measures of wealth for this community but would not be surprised to find that African Americans there do not deviate much from the national pattern. This article distinguishes income from wealth — income being annual earnings while wealth is the retained earnings after expenses. While African Americans in Prince George’s County may have high incomes, the real question is how much of that income is retained as wealth after expenses.

Persons who have studied wealth trends have noted that the wealth gap between middle class whites and middle class blacks in spite of similar earnings, leaves blacks in a more precarious situation. The more precarious situation is explained by the lack of wealth to provide a security net for the family during economic hard times. In these times, white families are better able to negotiate these times without substantial diminishment of their wealth assets. Blacks, on the other hand, tend to lose their small wealth assets more quickly. As a result, blacks tend to be more susceptible to downward shifts in the markets recently occasioned by a recession.

III. The Problem

“Racism doesn’t just come dressed in white sheets or voiced by skinheads, but lies in institutions that, like the FHA, have quietly and often invisibly channeled America’s wealth, power, and status disproportionately to white people. Those advantages are passed on and accumulate, generation to generation, giving . . . [whites] a head start in life.”

A. The Advantages of Wealth

6Id. at ___.

7Id.

Wealth is more important than income. Those with wealth can afford better education, better access to health, live longer lives, experience increased security provided by a private safety net, and have peace of mind, more stability and greater mobility within society. Studies document that when blacks and whites earn the same income but there is still a gap in wealth, and black children perform less well academically than white children.\(^9\) However, where wealth is comparable in both the black household and the white household, black children perform as well or even better academically.\(^10\) In addition, when wealth and parent education are both controlled for, black children are more likely than white children to graduate from college.\(^11\) The wealth gap is not explained by merit, industriousness or achievement but racism and vestiges of slavery, jim crow, and modern forms of discrimination - de facto and de jure.

The largest wealth building asset is the home. Whites have specifically benefitted from racial preferences in the housing market. Segregated white neighborhoods saw equity appreciate at much steeper values than experienced in black neighborhoods. White flight sucked the equity out of those very same neighborhoods as African Americans moved in, paid the price and whites moved out, taking the equity with them. Furthermore, growing equity meant better schools, parks and libraries, private colleges and down payments on homes for white children. In declining black urban communities, job loss, decline in property values, and increases in property taxes to offset the decline in property values meant that black families had more expenses and less savings. Where white parents are in a position to help their adult children, the opposite is often the converse in black families – the adult children financially support the aging parents. When you control for income, whites with the same income as blacks and Latinos have more than twice the wealth. Overall, this disparity can be traced back to these early public policies that advantage whites over others in home ownership. The wealth gap is mostly due to home equity and family inheritance. This view of wealth building is vastly different from the prevailing societal views that the wealth is explained by merit represented by a hard work ethic and intelligence.

This racial gap was fueled in many ways by federal policies. Federal Housing Administration (“FHA”) low cost home loans offered in the mid 1930s through the 1960s discriminated against black home buyers while, at the same time, reducing the down payment amount and substantially increasing the time to pay off the mortgage for white home buyers.\(^12\) Also, communities where blacks were likely to purchase homes were deemed ineligible for the


\(^10\)Id.

\(^11\)Id.

\(^12\)Insert source
program if those communities had even one or two black families. These communities were ipso facto labeled financial risks, and therefore deemed ineligible for the program. During this program, the federal government backed $120 billion in home loans, with more than 98% going to whites. Of the 350,000 homes newly constructed in North California from 1946 to 1960, less than 100 went to black families. As a result of such public policies, today, whites have ten times the wealth of black families and nine times the wealth of Latino families.

B. The Wealth Gap

Racism is the major cause of the wealth gap. Those same federal policies and programs described above which discriminated against African Americans have, in turn, privileged whites in the race to acquire wealth. In essence, increased wealth for one group over another is a spoils of a racist system, which in this case was (and is) fueled by federal policies and programs, including tax policies and rules that favor one group over another.

IV. Wealth defined

“A low-riding Honda with tinted windows, clothes sporting FUBU and Phat Farm, and a little bling bling are all signs of wealth. Well, at least, that’s what the people in my urban neighborhood seem to tell me.”

A. What is wealth?

Wealth is cultural and financial. Wealth is a snapshot in time of the well-being of an individual, family or community. Various cultural communities define wealth differently as the
opening quote to this section suggests. A popular notion of wealth in African American communities is best characterized as non-financial, i.e., family, children and education. For many, wealth has a variety of meanings and measures. In some Native American communities, wealth is reflected by what you give away to others as captured by this traditional saying, “It is more blessed to give than to receive.”  One writer suggests that “wealth is defined by what a person, culture or society values.” For purposes of this discussion, I will restrict my discussion of wealth to financial definitions of wealth. In doing so, wealth is the excess of assets over liabilities, and it provides a measurement of financial well-being that rises and falls over time. The excess is referred to as your net worth.

Typical resources counted in official governmental measurements of wealth include, but are not limited to, interest-bearing accounts, stocks and bonds, mutual funds shares, and home equity. The largest concentration of wealth by asset type is home equity. In 2000, based on data gather through the U.S. Census Bureau, Survey of Income and Program Participation, home ownership accounted for 32.3 percent of household wealth. All other asset groups (e.g., interest-earning assets at financial institutions, motor vehicles, IRA accounts, rental property, business or profession and U.S. savings bonds) account for less than 10 percent individually for household wealth. Stocks and bonds account for about 15% of total household wealth. While wealth is just a measurement at a particular point in time, it draws a powerful picture of where and on what is America’s wealth concentrated. Extreme Wealth is concentrated in the top 1 or 2 % of the society. Middle America measures its wealth much more modestly. These measurements show a strong correlation between household income and wealth accumulation. In addition to the favorable treatment given on the sale of a home, person in the top earning


20 Id.


22 Id. at 4.

23 Id. at 6 (Figure 1).

24 Id. at 8 (Table C). Households in the top 5th quintile had a net worth of $185,000 while households in the bottom quintile had a net worth of $7,396 in 2000.

25 Id. Household in the 4th quintile had a net worth more than 60 percent less than the top 20 percent.
brackets received preferential tax treatment on employer sponsored savings plans and other retirement savings plans that the ordinary taxpayer cannot take advantage of. These comparisons by race are even more striking.

The median household net worth for whites in 2000 was $79,400, while blacks had on $7,500 and Latinos had only $9,750. In essence, Blacks net worth was ten time less than whites, and Latinos income was nine times less than white. When compared by races at the same income level, blacks in the lowest quintile had a net worth of $57, Latinos had a net worth of $500, and whites had a net worth of $24,000.\textsuperscript{26} The numbers tend to show that there is a positive relationship between overall net worth and income, race, gender and household (married or singe) and job status. Nonprofit programs that address any of the factors would likely impact net worth but the largest asset, which accounts for net worth, is the home. To get a home, you generally need a dependable income source, and as the data suggests, there is a high correlation between income level and overall net worth. Income and home ownership seem to go hand in hand.

Redefining the contours of wealth in communities of color in order to better understand how wealth is perceived and accumulated in those communities requires consideration of notions of community well-being and individual responsibility for the collective. African American communities have remained viable through extended kinship circles beyond the traditional nuclear family. Grandparents, their children, and their children’s children often live under one roof, or otherwise co-depend on each other for food, housing, monies, jobs, and other resources. This co-dependence takes many forms including pooling of assets in informal groups as savings and wealth accumulation vehicles. This practice of pooling funds occurs in many people of color communities, and is often overlooked in formal studies of wealth practices in these communities. For example, the Koreans, Jamaicans, and Chinese have documented practices wherein individuals make weekly contributions to a general fund and those funds are distributed to members of the circle on a frequent basis pursuant to the rules of the group. These monies are often used for down payments on homes, job training and education, and start-up business costs.\textsuperscript{27}

As I am discussing wealth in this paper, I rely on the majority group view that wealth is

\textsuperscript{26}Id. at 12. The data also shows that the majority of net worth held by blacks and Latinos is held in the form of a house and motor vehicles. Few of the dollars are represented by holdings in interest bearing financial assets, such as, interest bearing accounts or stocks and bonds, or retirement accounts, such as, IRAs, Keogh accounts, 401K, or thrift savings plans. Id.

\textsuperscript{27}See generally J. Neimbhard, Cooperatives and Wealth Accumulation: Preliminary Analysis, 92, No. 2 The American Economic Review 325 -329 (This article discusses the use of cooperatives in communities of color to accumulate wealth.).
the excess of assets over liabilities. This definition is the dominant measure of wealth taken in federal studies of wealth and is also what is being measured in the U.S. Census Bureau’s Survey of Income and Program

B. How was it (is it) acquired?

Wealth can be measured at the levels of individual, family (including extended family), community and race. At the level of extended family, wealth is represented in the kin group and to some extent the pool of resources available to that kin group to meet the needs as defined by the kin group. At the level of race, the pool of resources is represented by the cultural habits of the sub-units of kin groups as influenced by larger societal interactions around real needs for wealth and perceptions of wealth as reflected in the political and power relationships. At the levels of the extended family, community and race, wealth is acquired through the pooling of resources. A nonprofit organization is a pooling source for wealth that can be strategically redirected to improve overall wealth indicators for the two lowest quintiles. In the end, the aggregate of wealth can be increased for a greater number of Americans, resulting in an overall better quality of life.

I would argue that wealth is acquired both passively and actively. Earned Income is the most common form of income that leads to active wealth. Income generated through jobs or engaging in an active trade or business is earned income, and adequate earned income is a precondition for wealth.28 Passive wealth, on the other hand, is commonly thought of as income generated by investments in stocks, bonds and other securities or interest-bearing financial accounts. This is not the form of passive wealth that concerns me. The form of passive wealth I am most concerned with is a much more insidious kind of wealth accretion. This form of wealth accrues to members of a race grouping simply because of one’s membership in that grouping. The person has not earned the wealth through productive activity or through the investment of earned income. Instead, government sanctioned preferential policies tend to account for most of the passive wealth accretion in the United States today. As an example, wealth occurred for large numbers of white Americans in the 40's and 50's when race covenants and redlining were active practices of the racial majority and reinforced by the political institutions of the state. Whites benefitted in enormous ways through programs that excluded blacks from homesteading, down payment programs for home ownership, government funds for education, and the like.29

With the ball rolling, whites acquired a privileged status in relationship to wealth.


29Insert source
accumulation, that today, only continues to widen the wealth gap.\textsuperscript{30} It is hard to play catch up with a ball and chain around your leg. The legacy of state sanctioned discrimination is still playing itself out on the wealth playing field. Today, blacks own\textsuperscript{10} times less the wealth of whites, and Latinos own nine times less the wealth.\textsuperscript{31} The wealth gap has widened in the last twenty years. Some have made the case that there is a hidden cost to being African American in this society, and it is this legacy that is being passed on from generation to generation, both black and white.\textsuperscript{32}

C. What role (if any) did public policy play in wealth acquisition?

State action was at the core of the redistribution of wealth in the United States. In the 19\textsuperscript{th} century, Thomas Jefferson, a slave owner and the third U.S. President, authored the yeoman-farmer programs, which had as its core purpose the taking\textsuperscript{33} of Native American owned land and the giving of that land to whites.\textsuperscript{34} The status of being white entitled white citizens, through no effort of their own, to benefit from one of the earliest government sponsored entitlement/welfare programs.

Another major state sanctioned hand-out to white Americans was done under the Lincoln
presidency through the Homestead Act. Blacks and Native Americans were specifically excluded from this program of giving large blocks of land for no money to white Americans. Many of America’s wealthiest families today benefitted directly from this practice, and can trace their wealth back to the acquisition of their core wealth generating asset – land. Land is a scare commodity, and once the land is given out, there is little that blacks and Native Americans can do at the level of the collective to catch up to the top 1 or 2 % in America. So the modest goal of my proposals is to raise the level of the wealth index in the lower quintile, hoping to increase overall well-being in black families and other communities of color.

Beyond the land welfare programs of the 19th Century, the GI Bill and Veteran Loans for Home Buying in the 20th Century provided education and home ownership dollars to white veterans, and not black veterans. Education is the single largest correlated variable to an individual’s ability to amass wealth. For most Americans with wealth, the single asset that accounts for almost half of the wealth of white Americans is the home. Home values outpace inflation and retain value better than any other single asset. Among the poorest of Americans, with little wealth, there is little wealth that can be tracked to home ownership. So it would seem logical that any proposal to increase home ownership would have an indirect effect of the rate of wealth accretion, and absolute dollars

I am not so naive as to think that simply owning a home is going to do it for black Americans. Wealth is the excess of assets over liabilities. Predatory lending practices, the absence of mainstream financial institutions for urban communities with a high concentration of people of color, and the increasing presence of private non-bank alternatives to fill-in the gap (e.g., check cashing stores and loan centers that loan money at near usury rates) has left communities of color even more vulnerable to the legacy of wealth deprivation that has plagued them for so long.

IV. The Solution - How do we level the playing field?

---

35Id.

36Id.

37Id.

38Some suggest that, in today’s market, investment in the private market is fastest way to accrue wealth. (Need cite). While this may be the case, the unpredictability of the private market may be a reason to discourage individuals from pursuing this strategy as an individual. I am open to the idea of pooling assets in group funds to spread the risk across a large group. Perhaps, federally funded black funds are required to enlist state action to re-level the playing field that because of state action created the lopsided playing field in the first place.
A. By using public policy/state action to re-draw the playing field

My proposal is simply at first glance. I am suggesting that we use public policy to eliminate the problem in the same way public policy was used to create the problem. Why not? The public policies embedded in the law of tax exempt organizations are a fruitful area to begin with. Tax-exempt organizations exist to fill in the gap between governmental and private action, with the goal of impacting the well-being of the citizenry. By definition, the 501(c)(3) organization is public, often works among the target population (i.e., the poorest of Americans), and has the confidence of most of the public. Further, this area of law already has a policy framework in place.

[insert history on nonprofit organizations]

B. What kinds of public policies are needed to level the playing field?

To ask a nonprofit to view its core mission as the accumulation of wealth among individuals is profound. Most nonprofits particularly in African American communities are predisposed to the crisis form of organization - food, clothing and shelter. I am not saying that those types of nonprofits are not needed and valued in the community; I just think not enough are thinking about how to harness the power of the nonprofit to serve as an intermediary in wealth building and transference programs in black communities. In my notion, the nonprofit becomes the surrogate for the family that transfers wealth intergenerationally. The nonprofit becomes the vehicle that amasses the wealth and transfers the wealth. The 501(c)(3) charity is the preferred route for all the same reasons that it is the preferred route over foundations for doing the work of a charity. Less oversight, less regulation and reporting obligations are needed to keep the organization agile and responsive to market and sector forces. To bog down the nonprofit in the quagmire of the foundation rules would be to stifle the movement before it ever got off the ground.

At the radical end of public policies needed to level the playing field, I am proposing a Black Tax Credit. The credit would be the equivalent of the forty acres and a mule that was promised to newly freed Africans once held in bondage in America. The promise was never fulfilled. The Black Tax Credit would be a way of achieving some measure of reparations through the Tax Code. Next, the kind of tax policies needed to level the playing field produce concrete and measurable results in terms of asset accumulation, which can than be tracked to wealth accumulation.

30I am aware of current debates and testimony on the Capitol Hill about the flagrant abuses some nonprofits are taking with their status as nonprofits. That debate aside, the nonprofit is still the most viable vehicle to redistribute wealth.
Next, at the level of the tax exempt entity, the most radical of policies would be the redrafting of the “no private inurement” clause under section 501(c)(3) to permit in limited circumstances the creation and transfer of wealth to targeted populations amassed in the bottom percentages of the wealth index. This provision would be retooled to open the door to more aggressive and creative use of the 501(c)(3) public charity as a means for lessening the wealth gap in America. This is necessary to free up charities to begin thinking in strategic ways beyond the current use of IDAs and other savings programs to lessen the wealth gap. The most recent evidence on these programs suggest that the population most benefitting from the programs are not the bottom quintile but those who live at 200 Percent above the poverty level. Those at the lowest end of the poverty scale are not gaining on the wealth index.

In addition to redrafting the “no private inurement” clause, there is a need to add to the list of permitted purposes under 501(c)(3) wealth accumulation for the poor. Adding wealth accumulation for the poor to the list of permitted activities removes the pursuit of such goals from obscurity to the conscious forefront of charitable work. Also, there is also a need to rethink political action by nonprofits to permit nonprofits to more forcefully lobby and intervene in political activity for the purpose of pushing the needs of the poorest among us who oftentimes are voiceless in the debate and fail as individuals to get the attention of politicians who ultimately decide if the laws are changed, and how they are changed.

Next, nonprofits that generate wealth in the form of profits in entrepreneurial enterprises should not have to pay tax on those funds provided those funds are wholly redirected to the nonprofit’s exempt purpose activities. This proposal would require us to rethink the provision on unrelated business income and the clause on organization purpose, which threatens the tax-exempt status of entrepreneurial nonprofits. Broadening the source of funding available to nonprofits would reduce the current reliance of such organizations on a striking pool of funds. The philanthropic pie is ever-shrinking and the demand is growing. This is good public policy particularly where the work of the nonprofit is closing the wealth gap which is enigmatic of so many other social ills plaguing urban communities.

An example of a nonprofit working in an entrepreneurial way to raise needed funds is Newman’s Own. The company solely exists to make products for sale in the open market, and the distribution of those profit to public charities. This practice is threatening to its nonprofit status because it operates solely to make a profit with no related charitable activities. Many

---

40Need Cite

41B. Shore, Charities Change Roles by Turning a Profit, USA Today, March 26, 1996 (Community Wealth Ventures web site http://www.communitywealth.org/art_charities_change_roles.htm).

42Id.
nonprofits because of current IRS rules create for profit subsidiaries to protect its tax-exempt status. As a result, significant funds are needlessly diverted to the tax coffers. I am suggesting that the creation of the subsidiary is not a necessary step if the rules are modified to reflect today’s markets. While this proposal seems to cut through the very traditional notion that nonprofits should not compete with for profit businesses, it is also a recognition that for profit businesses are limited in what they can do to fund all the necessary nonprofit work. By permitting some kinds of nonprofits doing asset accumulation work to compete more directly, the for profit business stands to gain from a more financially healthful consumer population. Everybody stands to gain from this proposal.

The work of transforming nonprofits is not just policy but attitudinal. I am speaking of the attitudes of those who create the nonprofits and those that service the nonprofits. The shift I am referring to is away for the social services paradigm of maintenance and problem solving each crisis toward long term social development for developing and building capacities of individuals, families and communities. This shift takes courage and a willingness to push at the outer edges of what is permissible, and to advocate for changes in the law if the current set of rules proves too restrictive in this regard. Persons do not typically think of private aggressive market strategies and nonprofits. Particularly when such strategies are for the sole purpose of transferring wealth assets to individuals. One of the core 501(c)(3) charity rules is that there should be no private inurement so when the nonprofit generates large amounts of financial assets for the express purpose of doling out those assets to its constituencies (individuals). It would seem that the rule against private inurement is being violated. These shifts in tax-exempt policy are necessary to begin to affect change at the institution level.

My proposals add to the existing web of policies currently in use to promote Asset Development among the poor. This movement has bipartisan support. In the next section, I will highlight some of the existing programs and where necessary push for an expansion of those policies beyond the current proposals.

C. What are Asset Development Policies? At the individual/family level? And at the community level?

Currently used wealth tools include the Earned Income Tax Credit, and some would argue that it is the most effective federal anti-poverty policy to date. While 40 percent of American live at or below the poverty line, the Earned Income Credit is responsible each year for lifting approximately 5 million people above the poverty line each year. Other tools include

---

43 Need Cite

the Child Tax Credit (partially refundable), and the Child and Dependent Care Credit (nonrefundable). In addition to the asset building tools, communities of color are replete with wealth draining ventures, like check cash stores, predatory lending, and tax refund anticipation loans. The target population spent more than $1 billion in 2001 for tax services, misdirect the EITC funds away from the targeted population and into the hands of tax preparers. As a consequence, any initiative that is tax based has to include methods and opportunities for nonprofits to provide free or low-cost income tax preparation for the targeted population. Other tools, Lifelong Learning Accounts, Child Savings Accounts. These policies are directed at asset accumulation and multiplying the efforts of the traditionally disenfranchised group to amass wealth. Each of these assets link to the conceptual framework of capital formation and wealth creation.

[To be Picked Up Here]

1. Individual/family level

Savings accounts

Homes

Workforce
   Skill training
   Daycare services

2. Community (makes communities better places to live)

Physical - community center or housing association

Organizational - learning centers, neighborhood crime watch or tenant association

V. Some Examples

A. IDAs (Savings Accounts)

1998 - pilot project

---

45Cite article on Ameriquest settlement here.

46Zdenek, supra note 28.

47Id.
WORKING DRAFT AS OF 3/27/05

2003 - expanded to tax credits

2006 - Discuss new bill expanding as of February 2006

B. Home Ownership Accounts (for acquiring, maintaining, and renovating)

C. [Discuss any other ideas]

VI. Examples of the role the nonprofit can play?

[Use institutional theory here]

Education

Investments in Community Assets

Matching Funds for Savings and Home ownership

Advocacy

Job Readiness and Placement

Entrepreneurial Training

VI. Conclusion

Will any of this make a difference?

Can it be legislatively done on a scale large enough to make a difference?

Are asset policies a substitute or complement for other necessary strategies?

Do these policies reinforce a self-reinforcing or self-neglecting politic?
INTERNATIONAL TAX RELATIONS: THEORY AND IMPLICATIONS

By Diane Ring *

INTRODUCTION

Governments challenge each other’s tax rules, power, and conduct, just as they challenge all spheres of social, political and economic life. How do these conflicts play out on the international stage? Who is successful and why? These are serious, critical and relevant questions for which we have few answers. The dominant focus of international income tax literature over the decades has been analysis of substantive tax law and its implications. Receiving much less attention is how countries have come to agree on particular tax rules and practices – the international relations of international tax.

Certainly the vast majority of tax rules are “domestic;” they are enacted and enforced by a single state. However, that characterization belies the inherently international nature of the subject. Many income tax rules directly affect large numbers of nonresident taxpayers engaged in cross border transactions and indirectly affect other nations. As early as the 1920s, countries began negotiating over the content and contours of income tax rules. The outcome of such negotiations determines each country’s prospects for tax revenues and for investment. Countries, though, are not the only participants in international tax debates. 1 Taxpayers, especially multinational corporations, and tax professionals actively seek to understand, influence and shape international tax law and policy. The current landscape reveals an enormous array of mechanisms by which the international tax system is influenced and these mechanisms vary

---

* Associate Professor of Law, Boston College Law School.

1 “International tax” has no formal or specific definition. It generally refers to the subset of income tax rules of a country that govern: (1) the taxation of residents of the country earning income outside of the country, and (2) the taxation of nonresidents earning income inside that country. In addition, income tax treaties, which are bilateral agreements negotiated between countries, are included under the umbrella of international tax.

2 Unless otherwise specified, the use of the term “regime” or “system” in international tax connotes the generic and broad usage of these terms in international tax literature, and not the more specific (though often
tremendously in terms of size, formality, membership, participation, subject matter and output. Relatively little attention has been devoted to understanding how these forces shape international tax policy – the focus on the “tax” piece of international tax has been at the expense of the “international” aspect which requires attention to the unique and complex dynamics of multi-jurisdictional relations. Not only must countries construct domestic tax policy and rules, they must also navigate the intersections of other countries’ rules.

The time has come to shift, or perhaps expand, the scope of tax analysis to include examination of the complex dynamics of multi-jurisdictional relations. This claim prompts two questions: “Why now?” And “Isn’t this addressed elsewhere?” As to the first, the multilateral nature of international tax has long been understood. However, two important features of the tax world have changed: (1) the volume of cross border business, and (2) the structure and form of that business. While both elements are widely discussed, their importance in affecting the dynamics and interactions that generate, develop and modify the international tax regime have not been explored. The increasing volume (and increasing number of participants) magnifies the importance of good cross border tax policy, and good policy requires more than identifying the desired substantive rules. It demands an ability to get them implemented in a complex multi-jurisdictional world.

The failure to anticipate other countries’ views or the interactions of their rules with your own can have potentially dramatic effects, including loss of revenue (to other countries or to

contradictory) uses of the terms in international relations (“IR”) theory, which will be examined more specifically in Part I.C.


4 Discuss rise of MNCs and the role of technology, intangibles.

5 See supra note 3 discussing the use of the word “regime” in tax literature as compared to international relations literature.

6 Although existing rules may have identifiable and known weaknesses, the limited volume of transactions or higher transactions costs of the past have operated as a friction on potential abuse of the rules. As the volume increases and the frictions decrease, the backstop for abuse is weakened and the underlying rules must be confronted directly.

7 Related to implementation of substantive rules is the impact they will have on other countries and the responses that will generate.
double non-taxation\textsuperscript{8}), increased antipathy to international investment,\textsuperscript{9} fairness concerns, increased administrative costs (as countries try to pursue regimes that are increasingly unworkable), and misallocation of investment and commercial activity. To avoid these counterproductive clashes, countries must be able to reach agreement on tax issues that require coordination for successful implementation. This need for agreement is not new, but as the number of important intersections between countries’ rules increases, so to does the pressure to secure agreement.

This conclusion leads to the second question, “Isn’t the subject of how countries reach agreement addressed elsewhere?” The answer is yes, and no. Study of the interactions among multiple participants in a global setting is the subject of an independent field of inquiry, international relations (“IR”). Unfortunately, literature and analysis in the IR area rarely use taxation as a case study for the theories of international relations.\textsuperscript{10} Thus, a nuanced understanding of the application of IR theory in the international tax realm remains lacking. This paper aims to begin filling this gap by exploring how we can use IR analysis of the international tax system to improve decision making in the international tax arena. The theoretical and practical aspects of negotiating with a multiple sovereigns and others in an interactive political and economic environment require specific attention.\textsuperscript{11} This paper starts bridging IR theory and


\textsuperscript{10} Use of game theory analysis, a prominent part of some IR analysis, appears in several notable works on international taxation. See, e.g., Avi-Yonah, supra note __; Roin, supra note __; Robert Green [International Tax/Trade]; Tsilly Dagan [Role of Treaties]. However, none of these works focused on the general question of how to examine comprehensively the interaction of states in the sphere of international taxation and the development of agreement on a range of topics in a field.

\textsuperscript{11} The claim here is not that strategic issues of international tax have been ignored (for example consider the initial efforts to compose international committees for drafting treaties in the 1920s). Rather, the rich and extensive work in IR theory that has been pursued over the past decades has not been integrated into international tax analysis. See infra text accompanying note ___ and Michael J. Graetz & Michael M. O’Hear, \textit{The “Original Intent” of U.S. International Taxation}, 46 Duke L.J. 1021 (1997).
international tax law with the ultimate aim of determining how countries reach agreement in international tax and how that knowledge can improve tax policy and tax negotiations. To test the proposition that IR will enhance international tax analysis, this paper identifies the most promising analytical direction in IR theory and applies that theory to the most widely known example of international tax negotiations among countries: the development of a system to relieve double taxation. Using this classic case study, the paper argues that the additional perspective and insight of IR theory can enhance our understanding of an established tax story. Ideally, such knowledge allows us to predict and, more importantly, modify the contours of international tax relations based on a knowledge of IR theory and tax goals. The paper also outlines an agenda for further research on these issues with the expectation that extensive case study projects are the foundation of predictive guidance on international tax relations.

Part I begins by establishing the normative goals of the tax system and then evaluating IR theory for use in tax debates. Part II examines the application of IR theory to the double taxation case study to cast this well-known story in a new light. Through this process, the paper evaluates which models of IR theory seem most pertinent to the structure, issues and dynamics of international tax. Building on such conclusions, Part III proposes an IR-international tax research agenda that would develop the IR analytical approach and extend the inquiry into many important substantive and procedural issues in the international tax system. The task of exposing the link with IR and exploring it in detail brings some common intuitions, past experience, and perhaps misconceptions, to the fore in an effort to see the international tax system through IR theory. The Conclusion considers further interaction between the tax and IR fields, as well as limits on IR theory’s ability to improve tax policy. Not only is the tax world called upon to integrate important areas of nontax research into its universe, but the scholars and researchers in the area of IR theory are exhorted to take seriously the international experiences of the tax system, especially as non-militaristic aspects of international relations, including the economic aspects, form an increasingly central part of IR theory. A marriage of IR theory and international tax -- in research, analysis and application -- should advance both fields.
I. DEVELOPING AN IR THEORY FRAMEWORK FOR INTERNATIONAL TAX

A. Introduction

The ultimate quest of this project, as well as much of international tax literature, is to improve tax policy. This goal demands a vision of what constitutes “good” – and how that will be measured. It also requires flexibility in exploring avenues outside of traditional tax analyses that may clarify important operational aspects of the international tax system. This Part outlines these two fundamental components of successful tax policy (a substantive tax goal and a model of tax relations). First, and more briefly, given its relative ubiquity in the literature, the goals of the tax system are defined and probed to identify their essential parts and their ambiguous or contradictory elements. Second, IR theory is examined to identify the research and theoretical models most likely to inform analysis of the international tax system. Then, Part II tests the usefulness of the IR model in understanding success or failure in international tax by applying the IR theory to the double taxation case study.

B. Normative Goals of the International Tax System

The basic goals (in addition to the obvious need to generate revenue) of the international tax system are the same ones generally espoused for purely domestic tax policy: efficiency, equity, and administrability. The list is simple; the contextual application is quite difficult. In the domestic setting these criteria are challenging to implement. Indeed, the components can even be difficult to measure because of informational limits (e.g., in efficiency oriented analyses) or uncertainty in the normative dimension of the standard itself (what constitutes equitable tax treatment). Although in theory one can analyze the efficiency and equity effects of, for example, the earned income tax credit, the picture is incomplete without consideration of other tax provisions impacting children and the family as well as nontax provisions with similar effects.12 Even when tentative conclusions can be reached regarding the efficiency, equity or administrative effects of a given tax treatment, there remains the question of how to balance among the three criteria in judging the desirability of the specific treatment.

In moving to the international context, additional questions arise due to the multiplicity of

12 See, e.g., Kaplow [2004 draft of Taxation and Redistribution, at Chapter 7]
sovereign taxing authorities, many of which are interested in the same taxable transactions. What do efficiency and equity mean in the global context? How do they impact the likelihood of reaching agreement on international tax matters? These terms are challenging to define in the abstract, to measure in theory, and to use to predict behavior. Starting with efficiency, should the goal of international tax be worldwide efficiency (either capital export or capital import neutrality)? What if that clashes with national efficiency? The dominant approach in the economics literature has been to evaluate international tax rules and policy in terms of worldwide efficiency. This theoretical starting point seems quite distant from the real world behavior of nations, given the potential for conflict between world-wide efficiency and national interests. If the goal here is to understand and predict how nations reach agreement, then a willingness to anticipate and see the impact of a national perspective on actual negotiations is critical.


14 If a rule favors certain investors among the universe of those who might import capital into in a single market, then the rule violates capital import neutrality (“CIN”) – and thus distorts savings incentives, and may impact the competitiveness of business. For capital import neutrality to exist, all investments made in a particular country should be taxed at the same effective tax rate, regardless of the identity of the taxpayer. Essentially, the total tax burden borne by all taxpayers investing in the country should be the same, even though the taxpayers may be based in a wide range of countries. See Graetz, supra note __ at 272.


16 Even with that premise, however, there remains the task of measuring efficiency, in particular the choice between capital export neutrality (“CEN”) and capital import neutrality (“CEN”), although some analysts reject both CEN and CIN as potentially too narrow, and instead encouraging a broader vision of economic neutrality. See, e.g., Daniel J. Frisch, The Economics of International Tax Policy: Some Old and New Approaches, 47 TAX NOTES 581, 586-87 (1990). Unless the tax bases and the tax rates across countries are the same, it is not possible for a rule to satisfy both CEN and CIN. See Thomas Horst, A Note on the Optimal Taxation of International Income, 94 Q.J. Econ. 793 (1980); Graetz, supra note __ at 272. Moreover, despite the dominance of the worldwide approach to measuring efficiency (through the use of either CEN or CIN), some have advocated as least a rethinking of national efficiency and whether it has a valid place in the efficiency analysis. After all, tax is somewhat unusual in that it typically assumes little weight for national efficiency. In many nontax
With respect to equity, what should be made of “inter-nation” equity – a factor not present in the purely domestic analysis? What kinds of claims can nations make upon one another regarding the relative and net impact of tax rules on different nations? The question of inter-nation equity incorporates political and moral judgment. The development of the equity analysis domestically does not encounter this question because it operates (in theory) in a closed system with a single sovereign making decisions based on the collective needs of the members of the society. However, the domestic literature, particular on equity, may be valuable in its exploration of what constitutes the source and nature of commitments to equity. To the extent it proceeds from explicit participation and membership in a common political venture (a single state) or from a broader membership in society, the conclusions may differ in an international context. In any event, each taxing authority’s view of equity (just like each authority’s view of efficiency) may starkly differ if the authority is considering the welfare of its own state rather than adopting a broader global view.

Not only are the efficiency and equity questions difficult in and of themselves for countries in their quest to design tax policy, these factors also affect the third prong of tax policy – administrability and implementation. It is insufficient to simply determine the desirable tax rules or regime for a given international tax problem. Unless actually implemented, they are useless. While administrability is obviously important in the domestic tax context, it is critical contexts we expect that nations, including the United States, will measure options from their own point of view and pursue what they understand to be their interests. See Graetz, supra note __ at .

See, e.g., Avi-Yonah, supra note __ at..

In this context it is important to distinguish between “equity-like” arguments that really are efficiency claims in disguise – for example an argument that a given tax rule benefits Country A at the expense of Country B and thus is undesirable on equity grounds, may be bolstered by observations that if the tax rule is implemented, the likely repercussions (including avoidance behavior, fall-out from declining economic conditions) will negatively impact Country A. This latter argument is less about equity and more about providing a complete assessment of the benefits and costs of the proposed policy – an efficiency argument.

Except to the extent that one seeks to recast even these questions as part of an efficiency framework, which mutes their independent analytical and explanatory value.

This complexity in the equity calculation helps explain the findings discussed later in the paper regarding states’ interest in relative gains from an international arrangement. Specifically, in assessing options, states may care whether “everyone gains” or whether “they gain more than everyone else.” See infra text
and tenuous in international tax where execution of tax policy frequently requires agreement among states on the issue. Enactment of domestic legislation is often inadequate to achieve desired tax policy. The tax rules and decisions of other jurisdictions are crucial to the bottom line for taxpayers with cross border income and for countries seeking to tax them. Thus, to effectively implement a desired tax policy it may be necessary to persuade other countries to participate in a shared vision, at least to some degree. Under what conditions are countries likely to agree? In answering this question, the paper turns to IR theory and analysis. Utilizing models from that literature and research, we may make better and more informed decisions about achieving cooperation.

Certainly, all three basic goals of the tax system demand recognition of extra-national concerns. Even if a nation takes a parochial view of equity or efficiency, administrability will require international policy coordination (and this may in turn require a readjustment of one’s views on acceptable equity or efficiency results). This need to coordinate provides a powerful impetus to develop a viable application of IR to tax. However, the need is even greater than first appears. Tax policy is not made in an international vacuum. Other aspects of international relations may impede or facilitate discussions of the taxing authorities. International disputes in unrelated arenas may prevent negotiations or the consummation of agreements. Moreover, because tax ties directly into the economic engine that is central to international power and prestige, it may be challenging to obtain agreement.

accompanying note ___.

21 Agreement can take place on a variety of levels, e.g., bilateral treaty, model agreements.

22 In addition to the typical domestic range of administrability concerns here, such as the level of information needed to implement the rule.

23 See e.g., I.R.C. § 901(j) which denies a credit for foreign taxes paid to a country if: (1) the United States does not recognize the government, (2) the United States has “severed diplomatic relations with the country” (or simply does not conduct relations with the country); or (3) the country is designated by the Secretary of States as one which repeated provides supposed for acts of international terrorism.” Countries currently identified by the IRS as falling under § 901(j) are Cuba, Iran, North Korea, Sudan, and Syria. Rev. Rul. 2005-3, 2005-3 I.R.B. 334.

24 For example, consider the absence of treaties with countries such as those identified in note ___.

25 In some situations, nontax issues will be integrated into tax analysis at an early stage either as a
The following section identifies the domains of IR theory valuable in assessing agreement in the international tax setting. The relevant IR models are elaborated in some detail, in preparation for Part II, in which the historical case study of double taxation is examined. It is important to reiterate that expansion of tax analysis to include applications of IR theory will not miraculously eliminate the challenges of international tax. Many of the core and persistent problems of international tax derive from the sometimes intractable dilemmas of efficiency, equity, administrability and the competition for tax revenue. Nonetheless, any information that improves our ability to forecast and pursue agreements remains crucial.

C. Mining IR Theory to Improve International Tax Policy

1. General Landscape of IR Theory

The IR literature is rich and complex, addressing a wide range of important questions about the conduct of international relations including the ways in which agreement is reached (or not) at the international level. The task here is to locate that segment of IR literature asking the same questions as this paper – How and under what circumstances can agreement be reached internationally? The field of IR research directed at this question, “international regime theory,” seems a sensible starting point. The next section outlines in some detail the contours of international regime theory in an effort to draw from that literature a model for analyzing the international tax system. However, this section briefly reviews the foundations of IR theory to provide some context for the discussion of regime theory to follow.

At a broad theoretical level, two important traditions, neo-realism and neo-liberalism, drive IR analysis.26 The former emphasizes: (1) the state as the central (and rational) actor; (2) the balance of power; (3) the importance of survival in an anarchical society; and (4) the

---

factor in efficiency or equity, or through additional policy arguments raised in the midst of the tax debate itself. For example, states (and taxpayers) often urge that their sovereignty is being undermined. Although sovereignty is typically not defined, the term tends to invoke an image of harm to the ability of the nation to speak for itself and make decisions that further its national interests in the global sphere. The role and meaning of sovereignty in international relations has been extensively considered in IR, and an illumination of that concept should better inform consideration of such claims with regard to tax policy.

26 See, e.g. CONTENDING THEORIES OF INTERNATIONAL RELATIONS, James E. Dougherty & Robert L. Pfaltzgraff, Jr. (2001)
importance of structure in terms of the relationship among units of the international system.\(^{27}\) This thinking is perhaps most closely identified with the work of Kenneth Waltz.\(^{28}\) The latter category, neo-liberalism (sometimes referred to as neo-liberal institutionalism identified with Robert Keohane), also takes the state as a primary actor. However, it views the states’ pursuit of national self interest in a market-oriented model\(^{29}\) as a critical factor in the outcome of international relations and in how successful international institutions can be in directing and modifying international behaviors.\(^{30}\)

In the latter part of the 20\(^{th}\) century a pluralist approach to IR theory emerged encouraging examination of the role of individuals, bureaucracies, and nongovernmental organizations in decision making at the international level. Although the literature frequently characterized neorealism and neoliberalism as in direct opposition to each other; the work of the pluralists played a somewhat different role. Rather than seek to completely supplant either of the two main traditions, the pluralist approach contended that neorealism and neoliberalism failed to account for a major force in the international relations dynamic when they ignored the role of nonstate actors. Thus, the pluralists’ work emphasized actors, relationships, and other influential factors.\(^{31}\) Examples include research on cooperation and integration in the European Union\(^{32}\) and on geographic environment.\(^{33}\)

\(^{27}\) Stein, WHY NATIONS COOPERATE at 4 (defining realism).


\(^{29}\) Typically viewing regimes as a creation of states to overcome interstate “market failure.”

\(^{30}\) Stein, WHY NATIONS COOPERATE at 7 (Against the backdrop of international anarchy, “order emerges as self-interested actors co-existing in an anarchic environment reach autonomous and independent decisions that lead to mutually desirable cooperative outcomes.”)

\(^{31}\) Although the description of the relationships between and among neorealism, liberalism, and pluralism draws on widely shared positions, any effort to group IR theories according to their meta v. medium range scope remains open to debate and not universally agreed. See, e.g., CONTENDING THEORIES OF INTERNATIONAL RELATIONS, James E. Dougherty & Robert L. Pfaltzgraff, Jr. (2001) at 17-18.

\(^{32}\) Id. at 519. Referred to as neo-functionalism and associated with Hass, Keohane, Nye, Schmitter and Lindberg.

\(^{33}\) See, e.g., the work of Harold and Margaret Sprout.
A final, significant IR strand, cognitivism, arises from the post modern IR literature.\textsuperscript{34} The focus of the cognitivist theories is less on the developing of paradigms for analysis and more on the critiquing of the dominant models for their “failure” to question the basis for knowledge. In particular, cognitivists direct attention “to the identity of those who make the choices on which reality is socially constructed.”\textsuperscript{35} These theories frequently question the researcher’s ability to construct “a value-free objective, or completely unbiased social science.”\textsuperscript{36}

Thus, neorealism, neoliberalism, pluralism, and cognitivism are the major threads of the IR landscape against which regime theory developed. Certainly the reality of how IR theory has emerged over the decades is less neat, structured, and easily contained than the above outline suggests. Comprehensive historical overviews of IR theory and its development struggle to present a coherent yet accurate picture. That said, there are in fact two important and clearly dominant traditions, and there have been two valuable additions to the thinking of those traditions – one more of a potential complement (pluralism) and the other a check and critique (cognitivism). Having established the role of these primary forces in broader IR theory, we can better understand the development and content of the more topical theory,\textsuperscript{37} regime theory, bearing directly on our tax question of when and how agreement is reached in international tax. The value and importance of the ideas imbedded in all of these IR threads surface in regime theory – adding both to complexity and to a useful richness. Much work in regime theory synthesizes the neorealist and neoliberal traditions, drawing upon their common features and shared assumptions (such as focus on states as central, rational actors and the reliance on a rationalist mode of explanation that emphasizes material factors, causal connections and

\textsuperscript{34} Post modern is an umbrella term covering diverse theories such as critical theory, feminism, constructivism, and post-structuralism. See Dougherty, \textit{supra} note __ at 38.

\textsuperscript{35} Dougherty, \textit{supra} note __ at 39.

\textsuperscript{36} Dougherty, \textit{supra} note __ at 38-39.

\textsuperscript{37} Regime theory does not purport to be a comprehensive “meta” theory widely governing all aspects of international relations. Rather it is a more topical inquiry – a theory that seeks to explain a smaller level phenomenon, the creation of an international regime.
scientific inquiry).  

2. **International Regime Theory**

   a. *Introduction:* International regime theory (“regime theory”) is that part of the IR literature trying to answer the question of how and when countries reach agreement – how they cooperate.  

   “Regime theory [seeks] to explain the possibility, conditions of consequences of international governance beyond anarchy and short of supranational government in a given issue area.”

   Although some basic version of these questions have always been of interest to students of political and international relations, attention to the idea of regimes developed in the 1970s.

   Several background factors prompted this interest, including a recognition among American scholars that the United States’ uniquely powerful post World War II status had enabled it to help create international regimes, and yet that power (as least in economic terms) was diminishing. This apparent change in dominant status generated questions about the prospects for future international agreements and regimes. What began as a research agenda among American scholars transformed into a long-lived and more encompassing inquiry joined by an increasingly diverse group of researchers.

   Regime theory is not a single theory in the sense of a single explanatory variable for the formation (or the content or effects) of regimes. In fact, part of its appeal is its reliance on the major meta-theories (neorealism and neoliberalism) in a way that intuitively seems plausible and enhances regime theory. Most scholars reject a single variable approach to regimes. Instead,

---

38 For example, the works of Barry Buzan, Charles Jones and Richard Little contribute to this effort. See Dougherty at 40, 88. A synthesis is plausible given the common “core assumptions” shared by realism and liberalism. Stein, *supra* note at 8. See *infra* text accompanying note __./.

39 Although most of the literature has a statist orientation, there is also emerging consideration of international regimes among non-state players (e.g., members of an industry). See, e.g. list.

40 RIT at 392-93.

41 See, e.g., Ruggie.,


43 RIT see chapters talking about the German-U.S. comparison and about a EU focus to the research.
they urge the development of a cohesive multi-variate theory of regimes, with an expectation that
the application of regime theory and the strength of certain predictive factors may depend on the
subject matter.\textsuperscript{44}

One certainty we have is that nations, taxpayers, organizations and others are now
participating in an endeavor loosely referred to as the international tax system and any systematic
effort to shed light on this arena can be helpful in a number of ways. First, the work of tax
scholars examining international tax using IR theory may encourage IR researchers to devote
more time and resources to this application area (defense/military and environmental examples
traditionally dominate the databases of IR researchers). Although this general work in regime
theory is applicable to tax, there are refinements that may vary according to the special features
of the topic area. Thus, specific examination of international taxation is desirable. Second, the
questions and ideas originating in IR theory may encourage different ways of thinking about the
players, the behaviors, and the role of international tax. Third, once the preliminary connection
is made between IR and international tax (here pursued through the question of regime
formation), it should become apparent that other major questions of IR theory and analysis will
be valuable to tax as well. Part III outlines a recommended research agenda to maximize the
contributions to international tax from regime theory.

b. Definition of Regime: What is a regime (or more precisely, an international regime)?
A “consensus definition,”\textsuperscript{45} drawn from the work of Stephen Krasner, has formed the basis of
much work in regime theory and served as a central thread. Krasner defines regimes as “implicit
or explicit principles, norms, rules, and decision-making procedures around which actors’
expectations converge in a given area of international relations. Principles are beliefs of fact,
causation, and rectitude. Norms are standards of behavior defined in terms of rights and
obligations. Rules are specific prescriptions or proscriptions for action. Decision-making

\textsuperscript{44} See generally RIT at 413.

\textsuperscript{45} Andreas Hasenclever, Peter Mayer, & Volker Rittberger, \textit{Theories of International Regimes}
(1997) at 8; see also RIT p. xii; RIT at 23 [Keohane quoting Krasner definition], RIT p.96-97 quoting and working
with Krasner definition], Richard Little, “International Regimes”, in The Globalization of World Politics (John
Baylis and Steve Smith, ed. 2001) at 303; Kenneth Abbott, “Modern International Relations Theory: A Prospectus
procedures are prevailing practices for making and implementing collective choice.”46 A modification of the Krasner definition, which has gained wide support (and in fact could be viewed as a clarification of an implicit assumption) is the emphasis on effectiveness.47 Thus, for example, a regime would not exist where certain rules have been nominally adopted by countries, but in practice are generally disregarded. Even if true effectiveness is not required, it is anticipated that the rules of the regime be “referred to in an affirmative manner by governments.”48

A further interpretation of the “regime” definition contends that “[o]nly where compliance is inconvenient, that is, where regime rules conflict with governments’ perceptions of what their self-interests would be if there were no such institutions – is the impact of the regime tested.”49 The strength of the inconvenience required should not be overstated. For example, aviation standards regarding common air traffic controller language (adopted to facilitate international travel and safety) result in many countries requiring airline/air traffic controller personnel to speak a language other than the native language of that state.50 Such a commitment may be inconvenient for the participating country, and thus conform to the regime concept, although the inconvenience hardly rises to the level of, for example, an interest in refraining from certain polluting or military activities.

46 Stephen Krasner, “Structural Causes and Regime Consequences: Regimes as Intervening Variables,” in INTERNATIONAL REGIMES (S. Krasner ed. 1983) at 2. Krasner’s definition has been valued in part because its complexity encourages greater specificity in analysis of regimes. See, e.g., Hasenclever et al., supra note __ at 12-13. However, another more simplified definition (which for our purposes is not in opposition to the Krasner definition) comes from Keohane: “Regimes are institutions with explicit rules, agreed upon by governments, that pertain to particular sets of issues in international relations.” Robert Keohane, “Neoliberal Institutionalism: A Perspective on World Politics,” at 4 in INTERNATIONAL INSTITUTIONS AND STATE POWER: ESSAYS IN INTERNATIONAL RELATIONS, (Keohane, ed. 1989).

47 See, e.g., RIT at 9, 27.

48 RIT at 28.

49 RIT at a33.

An important element of the regime concept that emerges more explicitly from Keohane’s definition of regime is the idea that it pertains not to an entire field (e.g., defense or the environment) but to a narrower problem or question on which some level of agreement has been reached. Agreement may be reached on certain aspects of a field but not others.\(^{51}\) As explored in the next part, this observation holds for international tax as well. It may be useful at this stage to differentiate among international regimes, international institutions, and international organizations.\(^{52}\) The terms can in some contexts be used loosely and overlap, however, there are important distinctions which help illuminate activity undertaken in the international tax arena. The category, “international institutions,” serves as the broadest term.\(^{53}\) International regimes, as defined above, are really a subset of international institutions, as are international organizations. The latter refers to a formal body that can serve as an information builder, or can do much more in the regime process.

c. **Outline of Regime Theory:** The focus of regime theory, as indicated earlier, is to explore and better understand the interactions captured under the concept of an “international regime.” The inquiry, research and analysis are divided into three basic questions:\(^{54}\): (1) how are regimes formed\(^{55}\) (including success/failure, timing, and substantive content of the regime);\(^{56}\) (2) what explains the form and operation of different regimes; and (3) what are the effects of the regime.\(^{57}\) To date, the bulk of IR research has tackled the first question; although the other two are exceedingly important and likely to be interconnected (for example, different types of

\(^{51}\) Insert environmental examples.

\(^{52}\) See, e.g., U.N. National Security council. [expand from RIT.]

\(^{53}\) RIT.

\(^{54}\) RIT at 406 (“three main tasks of regime analysis as a theory-oriented endeavor: explaining the formation, persistence and demise of international regimes;” “accounting for regime properties and their changes;” “and finally, determining regime consequences (or effects) and explaining their variation.”

\(^{55}\) And correspondingly maintained or subject to collapse—although these experiences are not inherently mirrors of the factors and explication of regime formation.

\(^{56}\) RIT p. 224.

\(^{57}\) See, id. at 29, 32.
regimes may generate different effects\textsuperscript{58}).\textsuperscript{59}

This paper utilizes the work of IR theory on regime formation—asking the core question in this tax area, why is cooperation achieved on some matters and not others?\textsuperscript{60} The current regime theory literature can be grouped into the following rough categories: (1) neorealist based regime theories, (2) neoliberalist based regime theories, (3) cognitivist based regimes theories, and (4) synthesis approach to regime theory (which includes attention to nonstate actors as suggested by pluralist theories in IR).\textsuperscript{61} This section briefly reviews the four perspectives on regime theory and then considers how they might be adapted for international tax.

i. Neorealist-based regime theory

The neorealist research starts with the expectation that states are motivated by their interest in their relative positions of power (including political and economic power),\textsuperscript{62} and go on to “emphasize relative power capabilities as a central explanatory variable and stress states’ sensitivity to distributional aspects of cooperation and regimes.”\textsuperscript{63} Thus, the key factors are relative power and distribution of resources among states. In this sense power can be both a tool and a goal. Under earlier versions of neorealism regarding regimes (hegemonic stability theory),\textsuperscript{64} it was argued that states created regimes where their common interests included public goods because in such cases the states needed cooperation to further their interests. But the

\textsuperscript{58} See, e.g., id. at 345-46.

\textsuperscript{59} See, e.g., Little, supra note ___ at __; RIT at 315-16 (noting that little scholarly attention has been directed to the effects of regimes).

\textsuperscript{60} A regime must first have been formed, before it can be studied for its particular operation and its effects. Given the dearth of IR literature on international taxation, formation seems a sensible starting point for examining the international tax system. It is anticipated that future projects will consider the second and third questions of operation and effects.

\textsuperscript{61} Andreas Hasenclever, Peter Mayer & Volker Rittberger, THEORIES OF INTERNATIONAL REGIMES (1997) at __. [also cite environmental projects and case studies]

\textsuperscript{62} RIT at 58.

\textsuperscript{63} Hasenclever et al, supra note ___ at 84.

\textsuperscript{64} Id. at 85-95.
thesis further specified “that only the presence of an outstanding economic and political power which ha[d] the capacity (and the willingness) to lead [could] make the group of states who participate in the world economy a ‘privileged group’ by supplying and supporting the infrastructure” [i.e. rulemaking and rule enforcement] for the regime.65 This means that regimes are created and sustained by powerful actors (power determined in relation to the issue area at stake) and that regimes should “decline” when the founding hegemon declines and power shifts.66

The continued existence of many international regimes long after the decline of their purported hegemon forced a rethinking of the role of power in the regime analysis. Recent power-based regime research has highlighted the impact of power in regimes for which more than one Pareto optimal outcome is possible (e.g., game theory example of the Battle of the Sexes)67 and thus coordination is required to avoid an outcome deemed undesirable by all. A classic example of such coordination scenarios includes the example noted earlier of specifying the standard language of communication for international air travel between air traffic control towers and pilots.68 Individual states would likely prefer their own language be selected, but they prefer agreement on another language over no agreement, and once that agreement is in place they have no incentive to defect.69 Acknowledging this role for regimes, neorealists assert

65 Id. at 90
66 Id. at 90.
67 In the Battle of the Sexes strategic game, a husband and wife each would like to go out for the evening together, but prefer different activities (A and B). In terms of their outcome preferences, both spouses would prefer to be together (even if not in the activity of their choice) than to be alone but engaged in their chosen activity. Thus, there are two Pareto outcomes - being together engaged in activity A or activity B. In either scenario, although one spouse is happier than the other, no change could be made to improve the situation of one spouse without making the other spouse worse off. See, e.g., Drew Fudenberg and Jean Tirole, GAME THEORY (1992) at 18-20.
68 See supra text accompanying note __.
69 Defection, i.e. choosing a language different from the other states confers no benefit because the harm from conflicting languages (now a real risk) exceeds the value of using your own language. See, e.g., Little supra note __ at 311-12.
that the important question to consider is why a particular agreement was reached.\footnote{Hasenclever et al, supra note __ at 104.} In addition, realists do not take the structure of the game as a given and argue that power may be employed to determine who is permitted to “play” (i.e., to the extent game theory models of regimes are used) and what rules govern the game process.\footnote{Id. at 104.}

Throughout this work, neorealists continue to emphasize states’ interests in achieving relative gains (as opposed to absolute gains). The belief is that states care significantly about their relative stature and an agreement that provides gains to both (or all) may still be unattractive if one state gains more than another (presumably because the state gaining more could secure some additional advantage by virtue of its relative gains).\footnote{Id. at __.} Neorealists do not view the attention to relative gains as an absolute determinant to regime failure – and are exploring the related factors that make a relative gain situation more or less acceptable.\footnote{Id. at __.} Moreover, though neorealists have traditionally accorded a smaller role to international regimes, they do not view them as irrelevant and, in fact, have developed a richer understanding of the possible roles: “[R]egimes may be used to affect . . . the variables that determine the severity of relative gains concerns”\footnote{Id. at 122 quoting Grieco.} that states face by: (1) giving the weaker party different treatment through the regimes; (2) facilitating side payments; and (3) granting the opportunity to voice views (especially the weaker parties).\footnote{Id. at 122.}

The neorealist understanding of states’ interaction, however, is challenged when power plays are less evident, either because no states seem sufficiently powerful in the issue area, or the powerful states are unclear as to what their interests really are.\footnote{RIT at 177-78.} Thus, room exists for

---

\footnote{Hasenclever et al, supra note __ at 104.}
\footnote{Id. at 104.}
\footnote{Id. at __.}
\footnote{Id. at __.}
\footnote{Id. at 122 quoting Grieco.}
\footnote{Id. at 122.}
\footnote{RIT at 177-78.}
alternative, or perhaps more accurately, complementary models such as neoliberalism, with explanatory power in these contexts.

ii. Neoliberalist-based regimes theories

The majority of regime theory analyses have developed under the umbrella of neoliberalism. Regime theories in this tradition start by envisioning “instrumentally rational actors”77 pursuing self interest and reciprocal benefits.78 Essentially, “the basic challenge for states is to overcome market failures. . . . [and] international regimes, one form of institution, are a device for overcoming market failure.”79 The idea of market failure here is that when the states compete with each other they may not reach their optimal outcome on certain issues because of some problems in the market structure. If instead of competing, the states act cooperatively they may achieve outcomes superior to the suboptimal outcomes of their competitive behavior. How can this cooperation be obtained? Through the formation of a regime that offers a number of solutions to the market failure, including overcoming the assurance problem, reducing transactions costs, and facilitating information exchange.80

Within this neoliberalist framework, regime research has pursued three major lines of inquiry, each targeting different factors in a regime formation: (1) the bargaining game involved,81 (2) the issue at stake,82 and (3) the background system.83

77 Id. at 56.

78 Id. at 55. This liberal strand has often been labeled “neoliberal institutionalism.” In this paper, the neoliberal position is generally identified by that general term (neoliberalism), although occasionally reference to particular authors draws upon the functionalism or institutionalism language).

79 RIT at __.

80 See, e.g., RIT at 35, 57, 80, ____.

81 As explored by work labeled contractualism/neofunctionalism and by “situation-structuralism.” See infra text accompanying notes ____. See also Arthur Stein, WHY NATIONS COOPERATE (1990) at 27-38 (describing coordination games, dilemma games and games of common aversion and the likelihood of success in each).

82 Seen in the work of situation-structuralism. See supra note ___.

83 The background system refers to both the impact of secondary variables (such as frequency of interaction among the parties) and the institutional bargaining involved. See infra text accompanying note
(1) **bargaining games**: The interactions among countries (and other actors) in the international arena can be viewed through the lens of game theory. For example, the game theory **prisoners’ dilemma** (“PD”) scenario\(^{84}\) illustrates the role of regimes in solving a “market failure” by providing a “collaborative outcome.” In a PD case, the parties will “defect,” even though that result is inefficient, because that outcome is the only equilibrium in the absence of “collaboration.”\(^{85}\) The creation of a regime facilitates the collaboration and information sharing necessary to overcome the parties’ fear of being cheated in a PD case. Extinguishing the fear of being cheated may be achieved through: (1) a monitoring function; (2) the implications of an iterative game; (3) the risk of tit-for-tat behavior; and (4) reputational concerns.\(^{86}\)

Of course there are an array of strategic games in addition to PD (coordination,\(^{87}\) .

---

\(^{84}\) The classic prisoner’s dilemma involves two individuals who have committed crimes and are now in custody. If neither prisoner confesses, there is otherwise insufficient evidence and they will both receive a light sentence. If one of them confesses, that prisoner is freed but the other prisoner receives a heavy sentence. If both prisoners confess they get a medium sentence. Clearly, the best result for both prisoners is to remain silent and take the light sentence. However, if each prisoner fears that the other will confess, it is better to confess and take a medium sentence (assuming both confess) than to be silent and receive the heavy sentence. Thus, without a mechanism to ensure their mutual silence (a “collaboration” that produces the pareto efficient outcome), they will “defect” (i.e. confess) which is a less desirable outcome, in order to avoid the worst possible outcome (not confessing while the other prisoner does “defect” and confess). See, e.g., Hasenclever, supra note \(\_\) at 30-31. A pareto efficient outcome is one for which there is no other outcome that both improves the position of at least one party and does not leave one of the parties worse off. See, e.g., id. at 45 n.18.

\(^{85}\) “Collaboration” and “coordination” are both distinct forms of “cooperation,” i.e., acting “together in order to achieve a mutually acceptable outcome.” Little, supra note \(\_\) at 314. Coordination, as seen in the air traffic controller example above, describes cooperation “to pursue a common strategy in order to avoid the mutually undesirable outcome arising from the pursuit of divergent strategies.” Id. Collaboration describes the cooperation in resolving a scenario such as PD, where the goal is to prevent defection “from a mutually desirable strategy [to] an individually preferable strategy.” Id.

\(^{86}\) Hasenclever, supra note \(\_\) at 35.

\(^{87}\) A coordination game differs from the PD collaboration version in that there is no risk of defection once some agreement has been reached. Close monitoring to ensure compliance is not necessary as it may be in PD case. For example, recall the air traffic controller situation. A single language for air traffic communication is necessary, and selection of a single language is the best outcome for all parties. Once that language is selected, no
suasion, assurance with different properties such as more than one Pareto efficient equilibrium (e.g., coordination games such as air traffic controller language), or alternatively, a single equilibrium where only one state is happy but the loser is still better off because defection is worse (suasion game). Depending on the type of strategic game characterizing the issue area, regime formation may be more or less likely. A hierarchy of probability for regime formation among strategic games has been articulated with suasion games being the most difficult in which to form a regime because the regime must not only address monitoring and sanctioning, but also distribution of goods (because the parties do not share the same dominant strategy. PD scenarios may be more capable of generating cooperation because distribution is not a major issue (the desired outcome of mutual silence imposes the same burdens on both parties) although monitoring and sanctioning needs are still present (recall that each prisoner needs to be certain that the other prisoner is not confessing). Coordination games are arguably easier still because only distribution needs to be addressed by the regime (e.g., picking the “winning” air traffic controller language). Once the selection (e.g., a language) is made, no country has an incentive to defect from that choice. Finally, assurance games are the most likely to experience success because the required coordination poses none of the three problems identified above (monitoring, one has an incentive to defect, nor does one really fear defection on the other side (as was the case with the prisoner’s dilemma which required some type of monitoring or enforcement of the agreement). In a coordination context, the agreement is “self-enforcing.” Hasenclever, supra note __ at 48.

The central idea in a suasion game is that one of the two parties has a primary strategy to cooperate, the “cooperation version” (or alternatively to defect, the “defection” version). In the cooperation version of the suasion game, the dominant strategy for party A is to cooperate –thus regardless of whether party B decides to cooperate or defect, A will cooperate. As a result, B has to be persuaded to cooperate because for B the best outcome is A cooperates and B defects, and this can be achieved automatically because A’s dominant strategy is to cooperate. Essentially, A has two choices: (1) threaten to act irrationally (i.e. threaten to defect even though that is against A’s best interest), or (2) offer something extra to B if B cooperates on this matter. It can be difficult, though not impossible, for A to persuade B that is will act against its own interests. However, it is possible to offer B something else for cooperation. Thus, we might expect to see a suasion game on one issue linked with another suasion game between the parties. See Hasenclever, supra note __ at 51-52.

The classic example of an assurance game is the stag hunt in which two hunters plan to go hunting for a stag. In order to hunt the stag they must work together; if either hunter wanders off to catch a rabbit then that hunter will gain a rabbit, but no stag will be caught. Thus, both hunters prefer to cooperate and hunt the stag. However, mutual defection is also an equilibrium because if one hunter defects, then the other should as well (both obtain rabbits). The worst outcome is if one hunter defects (goes after the rabbit) and the other hunter continues hunting for the stag alone (and thus catches no stag).
sanctioning, and distribution). In the assurance game, both parties share the same preference (joint cooperation), so a regime is needed only to “assure” each other that they have rationally reviewed the situation and recognize their mutually preferred outcome.

(2) issues: The issue area itself also constitutes an important explanatory factor for regime creation. Three major policy domains dominate international life: security, system of rule, and economic well-being. Security issues, quite obviously, concern physical protection from internal and external threats. System of rule issues are those regarding how states allocate “opportunities for exercising freedom and for political participation . . . among individuals.” And finally, issues of economic well-being address how economic gains are shared and how the “opportunities for achieving such gains” are allocated. Why does it matter to which category and issue belongs? The predictive conclusion is that if an issue area belongs to the economic category (where there is “divisible ‘gain,’ rather than indivisible ‘power’ . . . at stake”) then regime formation is more likely than for security or system of rule issues. As between security and rule issues, security is considered more conducive to regime formation.

In addition to this basic categorization of issues by subject area, conflicts (or issues) can be grouped based on the nature of the conflict itself. Conflicts break down into four categories: (1) “conflicts about means” – states share the same goal but disagree on how best to achieve it; (2) “conflicts about values” – states “hold incompatible principled beliefs regarding the

90 Hasenclever, supra note __ at 54.

91 “Issue-area” has been defined as “one or more, in the perception of the actors, inseparably connected objects of contention and of the behavior directed to them.” Id. at 60-61.

92 Id. at 62 (citing Czempiel).

93 Id. at 62 (citing Czempiel).

94 Id. at 62 (quoting Czempiel).

95 Id. at 62.

96 Id. at 62-63. Note that a classic example of a “rule” issue is human rights. See, e.g., Id. at 63.
legitimacy of a given action or practice;” (3) conflicts of interest over “relatively assessed goods” – states value the same scarce resource, in this cases goods whose value to the state depends on how much of the good the state has relative to other states (e.g., weapons); and (4) conflicts of interest over “absolutely assessed goods” – states value the same scarce resource, in this case goods whose value to the state is independent of the amount of the good possessed by other states (e.g., food items). 98

The first two classes of conflicts are “dissensual,” that is the states disagree about what is “desirable.” 99 The second two classes are consensual; there is agreement as to what is desirable and that agreement generates the conflict because the good is subject to competition (weapons or food, in the above examples). 100 Again, does it matter which type of conflict countries are facing? Are some more amenable to agreement than others? Regime theory work contends that conflicts over absolutely assessed goods (e.g., food) are most conducive to regime formation, followed by conflicts over means, then conflicts over relatively assessed goods, and finally the least conducive would be conflict over values (e.g., human rights). 101 The grouping of cases by type of conflict may strongly mirror the grouping by type of issue: security conflicts are likely to be conflicts over relatively assessed goods; economic conflicts are likely to be conflicts over absolutely assessed goods, and rule conflicts are likely to be conflicts over values. The remaining conflict type, “conflict over means,” reflects disagreement about form not substance. (Recall that this category is viewed as the second most likely to produce agreement). Thus, we would expect to see conflicts over means arise in all three subject areas (security, rule of law, and economics).

(3) background system: Features of the background system against which the specific

---

97 Id. at 63.
98 Id. at 63-64.
99 Id. at 63-64; Beth A. Simmons & Lisa Martin, International Organizations and Institutions, in HANDBOOK OF INTERNATIONAL RELATIONS (ed. Walter Carlsnaes, Thomas Risse, and Beth Simmons (2003) at 196.
100 Simmons & Martin, supra note __ at 196; Hasenclever, supra note __ at 63-64.
101 Hasenclever, supra note __ at 64.
international conflict is taking place can impact that likelihood of agreement and regime formation. The two major aspects of the backgrounds system are: (a) the set of “secondary variables” that characterize the parties and the situation, and (b) institutional features that drive the basic mindset of a country in the negotiation. The secondary variables can be readily enumerated and the potential influence on interaction between and among states is obvious, even if the specific outcome is not. These variables include: (1) “frequency of interaction,” (2) type of foreign policy practiced by the states, (3) “the distribution of issue-specific resources, (4) “the presence or absence of salient solutions, and (5) “the number of actors in the issue-area.” The effect of each factor may vary depending on the strategic game under consideration. For example, distribution of the relevant resources among states “makes a difference for coordination, collaboration and suasion situations, but hardly so for assurance situations.”

The “institutional bargaining” factors are elements inclined to affect the state’s general perspective on the agreement process. Two key elements are likely to promote bargaining: (1) uncertainty – states’ uncertainty about possible strategies and outcomes and about their relation to the states’ “core” interests leads states to “engage in integrative (rather than distributive) bargaining,” and (2) exogenous shocks – the existence of an exogenous shock (e.g., discovery of the ozone hole over the Antarctic) is expected to increase states’ efforts to create a regime. Although the “veil of uncertainty” facilitates regime formation, additional factors (similar to the background factors listed above) enhance the likelihood of success, including: (1) an equitable solution, (2) an effective compliance mechanism, and (3) strong individual leadership (relying on “negotiation skill and ingenuity (rather than power)).

102 Id. at 54-55.
103 Id. at 55.
104 Id. at 55.
105 Id. at 72.
106 Id. at 75.
107 Young; Hasenclever, supra note __ at 72.
Overview of Neoliberalism: In sum, neoliberal regime theories view regimes as forming in response to market failure in the interaction among states. Market failure itself though is insufficient to ensure the creation of a regime. Three core factors influence whether a regime will in fact emerge: (1) the bargaining game involved; (2) the issue itself; and, (3) the background system (institutional bargaining features and the “secondary variables”). Evaluating the bargaining game, issue, and background system of a particular international conflict may indicate the likelihood of regime formation and may reveal the greatest hurdles.

Despite the distinctions between neorealist and neoliberal theories reviewed here, the two are strongly connected as rationalistic theories and in fact are potentially quite complementary in their views of regimes. Both share a number of common assumptions, including: (1) states function against a backdrop that is anarchical, (2) states are rational (and, at least to some degree, unitary actors), (3) cooperation in the international environment is the source of regimes, and (4) regimes “promote international order.” These common features are the subject of critique from the major remaining strand of IR theory on the study of regimes, cognitivism.

iii. Cognitivist-based regimes theories

Cognitive theory stands separate from the first two strands of regime theory. Cognitive scholars as a group have provided the primary challenge to the rationalistic theories, “criticiz[ing] realists and utilitarians [i.e. neoliberals] for not taking into account the pervasive ambiguity of reality and consequently emphasiz[ing] factors such as perception, knowledge, and

---

108 See supra text accompanying note __. See also, RIT at 260-266 (describing the differing effects of conflicts over means, value, and goods (both absolutely assessed and relatively assessed).

109 Keys to successful regime formation include the nature of the available solutions and the presence of leadership. See supra text accompanying note __.


111 RIT at 30.

112 Richard Little, International Regimes in THE GLOBALIZATION OF WORLD POLITICS (John Baylis & Steve Smith ed., 2nd edition 2001) at 301. See also, [from RIT].
ideology.” In fact, the theory is predominantly a critical one, offering less of an affirmative vision of regime formation. Cognitivist theory in the regime area has divided into two strands, a “weak” version and a “strong” version,” although only the former (the “weak” version) contributes to regime theory. Under the weak version, “the demand for regimes in international relations depends on actors’ perceptions of international problems, which is, in part, produced by their causal and normative beliefs.” Thus, the weak version draws attention to the question “what are the ‘origins and dynamics of social actors’ understanding of the world?’” The analysis of regimes from the weak cognitivist viewpoint starts with three assumptions: (1) states’ interests are contingent on their understanding of the world; that is, these interests are not a given; (2) because of the technical nature of many international issues and the uncertainty that states experience in trying to assess what options best reflect their interests, states are increasingly relying on experts and scientists to provide information and advice (and such

---

113 RIT at 409.

114 Hasenclever, supra note ___ at 137. Note that the terminology of “weak” and “strong” refers to the strictness and exclusiveness of the theory’s focus on knowledge, not to the quality of the arguments.

Strong cognitivism is more radical and less capable of being unified in a multi-level analysis with the rationalist theories. (Major theorists writing in the area of strong cognitivism include Friedrich Kratochwil, John Gerard Ruggie, and Robert Cox). The strong cognitivists reject at the outset the rationalists’ view of states as calculating goal maximizers, and offer an alternative view of states as “role players.” Hasenclever, supra note ___ at 155. Given the limited usefulness of the strong cognitivists’ approach for this paper’s inquiry into regime theory, only a brief explanation is presented here: Under the “role playing” view, the states do not decide their actions based on the expected consequences (as they would in a utility maximizing view), but rather they decide how to act based on what is appropriate for an actor in that role. Therefore their behavior is described as rule driven. The state asks, What kind of situation is this? Who am I in this situation (i.e. what is my role)? And finally, What actions would be appropriate for a party in this role? See Hasenclever, supra note ___ at 155-157 (discussing the work of James March and Johan Olsen).

Some strong cognitivists also reject the underlying methodology of the rationalists: positivism and reliance on “objective knowledge” and the “formulation and empirical testing of causal hypotheses.” Hasenclever, supra note ___ at 220 (citing Keohane in Rittberger at page 24-26.). This paper, while cognizant of the limits of a positivist methodology, concludes that rationalist approaches combined with weak cognitivism can provide a more relevant and instructive model of international tax relations. Thus, Parts II and III proceed by analyzing the international tax system according to rationalist/weak cognitivist models – both individually and as part of a more integrated approach.

115 Hasenclever, supra note ___ at 137.

116 Id. at 137.
reliance affords the experts an opportunity to influence state action); and (3) some degree of
intersubjectively shared meaning about the subject at issue is required before regime formation is
possible; that is, the states must have some shared vision of the problem and its context in order
to achieve some level of cooperation.\textsuperscript{117}

One particularly interesting development from the work is attention to the role of learning
and “epistemic communities.” Because of the weak cognitivists’ first assumption (a state’s
interests are not a given), policy makers turn to experts as described in the second assumption, to
help them formulate the state’s interests. In some cases, these experts as a group may be able to
shape the direction of policy and, through widespread dissemination of particular knowledge,
facilitate broad agreement and ultimately regime formation. Such a group of experts is defined
as an “epistemic community” – “network[s] of professionals with recognized knowledge within
that domain or issue-area.”\textsuperscript{118} Epistemic communities facilitate regime formation by first
developing some consensus among themselves on an issue (e.g., ozone). Second, the epistemic
community, which exists across the relevant states, exerts influence to shape, direct and change
state views on the issue. This new knowledge and learning on the part of states can cause the
states to redefine their conception of their national interest.\textsuperscript{119} To the extent the epistemic
community has been successful disseminating its information and causing decision makers to
adopt its views, the “widely shared ideas may facilitate cooperation in the absence of a unique
equilibrium, [and serve] as focal points which help define acceptable solutions to collective
action problems.”\textsuperscript{120} Thus, weak cognitivism proposes a model of how states may come to adopt

\textsuperscript{117} The cognitivists’ focus on this last point does not make it exclusively their domain. Presumably
neorealists and neoliberalists anticipate that countries must have some measure of agreement on what the issue is in
order to work through a resolution. The distinction is that the cognitivists are particularly focused on why we know
what we know. If information shapes states’ goals and conduct, what is the origin of the information? What kind of
power and influence accords to those who possess, frame and disseminate information?

\textsuperscript{118} Peter Hass, “Introduction: Epistemic Communities and International Policy Coordination,” in
KNOWLEDGE, POWER AND INTERNATIONAL POLICY COORDINATION (special issue of International Organization 46,
no. 1 1992).

\textsuperscript{119} Cite re role of Ep. Com.

\textsuperscript{120} Hasenclever, \textit{supra} note \_ at 144.
new positions and how purveyors of knowledge can under certain circumstances shape the
direction of new policy. Despite cognitivists’ major role as skeptic and critics, cognitive theories
are frequently understood (even by its advocates) as being “complementary” to realism and
neoliberalism.\textsuperscript{121}

\begin{itemize}
  \item \textbf{iv. Synthesizing the regime theory literature}
\end{itemize}

For anyone trying to use IR theory as a tool to understand and ultimately improve a
substantive field such as international taxation, the question is how to envision the relationship
between or among the various IR theories. Are they purely competitive models requiring the
user to test and select the most appropriate one? Are they compatible and capable of integration
into a larger theoretical approach? Although much of the IR regime theory work has been
strongly competitive, with scholars pursuing their own individual directions, there have been
efforts to synthesize the IR literature. With respect to coordination of neorealism and
neoliberalism one could ask “whether market failure or distributional issues best describe the
range of issues involving international politics.”\textsuperscript{122} As an example, Krasner argues that “human
rights regimes . . . do not address market failure problems and therefore cannot be adequately
understood from the perspective of liberal co-operation theory . . . [i.e. neoliberal regime
theory].”\textsuperscript{123} He applies his position to four case studies of human rights and concludes that
“success and failure were not a function of the regime’s ability to reduce uncertainty and to
supply information for its members but rather of the willingness of the most powerful states to
enforce its principles and norms (which these states themselves had promoted) [i.e. neorealist
regime theory].”\textsuperscript{124} Whether we view international tax conflicts as the result of market failure
(e.g., uncertainty, information needs, and transaction costs) or not, may have direct bearing on

\textsuperscript{121} “Ijn my opinion, [cognitive theory] should be seen not as a substitute for, but as a complement to
existing regime theory.” RIT at 203. This move has been interpreted as not just a vote for complementarity, but a
recognition of the “priority ([but] not the superiority) of rationalistic approaches . . . .” RIT at 410 (Ritberger
assessing Jonsson’s position on the place of cognitive theory in the regime agenda).

\textsuperscript{122} RIT at 140.

\textsuperscript{123} RIT 408.

\textsuperscript{124} Id. at 408.
the applicability of neorealist v. neoliberal regime theories. It is certainly plausible, however, that different international tax issues could be classified in different ways. Even if market failure plays no role in any human rights regimes, such uniformity may not be evident in all other subject matter areas.

If regimes develop for different reasons, then the different theories may each be valid and informative for some subset of cases. How can we identify each subset of cases? Haas identifies four basic regime patterns:125

1. neorealist “follow-the-leader” pattern in which “[t]he regime is created by the strongest party and other countries are compelled by the dominant country to co-ordinate their policies” (although the regime itself may be weak or strong depending on the underlying goals of the powerful state);126

2. institutionalism and bargaining in which regimes are created through bargaining structures reflecting each state’s individual cost-benefit analysis and drawing upon game theory propositions (the more participants, the more difficult to reach agreement; moreover, the rule or position adopted by the regime often represents the lowest common denominator of the participants (although not always));127

3. epistemic communities and follow-the-leader which reflects the cognitivists claim about uncertainty in the international tax system – if states’ interests are unclear and epistemic communities have “consolidated influence in the dominant state, then [the] follow-the-leader [model] may be modified” to acknowledge the epistemic consensus and role in the hegemon;128 and,

4. epistemically informed bargaining in which a widespread epistemic community can lead to the creation of a regime even in the absence of a strong state; for example, “[e]nvironmental

---


126 Id. at 180-81.

127 RIT at 183.

128 Id. at 187. Essentially, the epistemic community in the hegemonic state is helping that state to formulate and identify its interests, which it will then, as “hegemon,” pursue.
regimes in this instance emerge through institutional bargaining as described by institutionalists.”

These four patterns represent a power-based model (realism), an interest-based model (liberalism), a modified power-based model where information or knowledge (cognitivist thread) impacts the exertion of power, and a modified interest-based model where the institutional bargaining pursued by the states is significantly impacted by information or knowledge transmitted by epistemic communities.

v. Adaption for International Tax:

The next Part undertakes the evaluation of the international tax case study using IR regime theory. The purpose is to assess the value of regime theory as an explanatory model. Given the novelty of reviewing international tax in these terms, an important goal is to determine what features of tax are more unique and may affect application of the theory.

First, we must discern the type of regime pattern best characterizing the international agreement under scrutiny. The dominant distinctions lie along the power/interest (i.e. realist/liberal) line. Second, we must analyze the case within the framework of either the neorealist or neoliberalist theory with attention to the knowledge-based questions generated by the weak cognitivists, where applicable. (Recall that of the four regime types, two are power based (with and without the cognitivist thread) and two are interest based (again, with and without the cognitivist thread).

Before continuing on to Part II and the analysis of the double taxation case study, it is valuable to confront an obvious reality of this project: Any effort to import IR learning into international tax requires a certain measure of effort on the part of the tax analyst. The IR structure, terminology, perspective and tools must be absorbed and understood before they can

129 Id. at 188. In this pattern, there are no relevant “hegemons,” and the states do not have a clear sense of their interests. If an epistemic community spread across the states reaches a degree of consensus on some issue, such as the pollution problem in a certain region, then the epistemic community may be able to mobilize in each of the affected states to inform decision makers (and the public) and galvanize their action.

130 Although many of the case studies completed by IR theorists to date tend to favor interest based theories as explaining regime formation, there remain gaps in that theory, notable in its failure to fully account for struggles market by power and distributive concerns. See, e.g., RIT at 229-30; Hasenclever et al., supra note __ at 4,
be applied to tax. Can we be sure that this investment is worth the effort? Several factors strongly argue yes. First, in the environmental law context, application of regime theory analysis to a series of case studies has generated interesting preliminary analysis about factors that are regularly present in cases of successful regime formation and those that are absent. For more than 15 years, environmental researchers and scholars have been studying regime formation and regime effectiveness.\textsuperscript{131} One large multinational collaborative project conducted five in-depth case studies in environmental regime formation.\textsuperscript{132} The project sought to ascertain the determinant factors in success (or failure) or regime formation. The five case studies were tested for the persuasiveness of four basic explanatory hypotheses (power based, interest based, knowledge based, and contextual analysis) – both broadly and in more detail. Interesting and potentially valuable observations emerged from this serious study. For example, the study suggested that knowledge based factors can play a role in early regime formation but are less significant as the terms of the regime are more concretely negotiated.\textsuperscript{133} “Leadership” of individuals (in particular, structural, entrepreneurial and intellectual leadership) played a notable role in all of the environmental case studies and was thought to be a “necessary condition” for the regime formation. Recognition of this environmental research\textsuperscript{134} (as well as other projects exploring regime questions including why regimes vary in their effectiveness)\textsuperscript{135} identifies a plausible role for such research to play in the development and growth of the substantive subject area.

Second, the effort to carefully parse the factors influencing international relations is likely to cause us to evaluate them more precisely in the international tax context, even in the absence of a full scale IR regime theory analysis. Without question, improving our understanding of how we do and do not reach agreement on international tax does matter.

\begin{itemize}
\item \textsuperscript{131} See, e.g., The Effectiveness of International Regimes,” edited by Oran Young (1999).
\item \textsuperscript{132} The project involved 14 researchers from four countries. See generally, “Testing Theories of Regime Formation,” by Oran Young and Gail Osherenko, in Regime Theory and International Relations, Rittberger, ed. (1993).
\item \textsuperscript{133} P. 246
\item \textsuperscript{134} For additional analysis of environmental regime formation, see “Creating Regimes – Arctic Accords and International Governance” by Oran Young (1998).
\item \textsuperscript{135} The Effectiveness of International Environmental Regimes, Oran Young ed. (1999).
\end{itemize}
Taking the time to consider these questions specifically should enable us to see new international tax questions in a different light.

Third, use of an independent field of study to enhance tax policy and practice is not new. For decades, in both domestic and international tax, economic analysis and economic models have played an important role in shaping discussion, targeting concerns, and framing possible options. As an example, in the debates surrounding domestic tax arbitrage, the economic analysis attempted to flesh out the impact of potential tax rules under different scenarios. Pursuit of economic analysis requires knowledge of, and attention to, the broader field of economic analysis. Even though the resulting economically derived work on tax rules typically relies on stylized models and highly simplified assumptions, there seems a general sense that the effort to integrate economic analysis into the development of tax law and policy has been, on balance, fruitful. The expectation is that international relations theory and analysis will further tax policy in the international arena.

II. INTERNATIONAL TAX CASE STUDIES IN REGIME THEORY

A. Introduction

The purpose of this project is to explore the relevance of IR theory, in particular, regime theory, to international tax. Clarification of tax terminology is therefore crucial before introducing the case study. Frequently tax practitioners, scholars and government officials speak of the “international tax system” or “international tax regime,” by which they have in mind a wide assortment of common rules, common problems, and interactions among nations. However, for purposes of exploring the ability to reach agreement on particular rules in the international tax arena, such a broad use of the term “regime” is not effective. Some features of the system constitute a regime but others do not. Describing the entire effort as a “regime” would be both inaccurate and would miss much of the detailed operations in the international tax world. Just as in the environmental or security area one would not speak (in the context of regime theory) of a single regime, but rather of regimes on different initiatives (e.g., the
Mediterranean Action Plan - targeting the poor water quality of the Mediterranean), so too in the tax context is it necessary to focus more clearly on particular problems or initiatives.

The case examined here, the emergence of the avoidance of double taxation regime ultimately embodied in the three major model income tax treaties (“double taxation regime”) is presented through an abbreviated history. It is fully anticipated that additional analysis could be pursued. Part III identifies some of these strands as part of its suggested research agenda.

B. Avoidance of Double Taxation

1. The Regime

The selected case study targets what is often characterized as the heart of the international tax system – agreements to relieve double taxation (i.e. income tax treaties). Although these agreements might be tempting primary candidates for regime designation and analysis, it is valuable to look behind the specific bilateral treaties to the network of bilateral treaties, the model treaties, and the domestic law which reflect a shared understanding of the importance of double taxation and its relief in the income tax system. Arguably a regime exists at this level, at least as measured by the consensus definition of regime: “the principles, norms, rules, and decision making procedures” that prescribe state behavior in an issue-area. Thus, in this case, the principle is that international double income taxation is harmful and should be avoided. The

136 RIT at 191-93.

137 The problem of double taxation refers to the possibility that two different states might tax the same income of a given taxpayer. For example, if a U.S. corporation earns money providing services in France, both the United States and France could plausibly and “validly” seek to tax that income. Such an outcome generally has been perceived as undesirable because, particularly at higher tax rates, such an uncoordinated tax burden could eliminate all profit from the transactions. (If the U.S. corporation faced a 60% income tax rate in both France and the United States on the same item of income, the total tax bill could exceed the income). However, relief from double taxation does not require agreement (bilateral or multi-lateral). A country can unilaterally provide this relief by offering a credit for foreign taxes paid (or alternatively by excluding income earned in another country from taxation). In fact, shortly after the U.S. income tax was enacted in 1913, the United States adopted a credit system allowing U.S. taxpayers a credit against their U.S. income taxes for foreign taxes paid on foreign income. Given this unilateral relief opportunity, actually utilized by a major economic player, what more was needed? Why was a regime needed? As noted infra certain issues required for the proper functioning of these credits depended on the cooperation of (and adoption of appropriate rules by) the corresponding source sovereign. Thus, unilateral action sought bilateral long term support. See infra text accompanying note __.

norm is that residence countries should yield primary tax jurisdiction to source, at least with respect to certain types of income. The rules include the details of the particular mechanisms by which the residence jurisdiction yields to the source jurisdiction (either through granting a foreign tax credit for the income taxed by the other country, or through the exemption of foreign source income). The procedures include the process of bilateral specification through a negotiated treaty with reciprocal rights and obligations and the opportunity for review through the competent authority mechanism.\textsuperscript{139} Although the details of many individual treaties vary, their content is remarkably constant, down to the order and numbering of the actual articles. This uniformity holds not only for the major model treaties (OECD, U.N., and U.S.), but also for the network of bilateral treaties. How did this regime emerge? What were its origins? These questions must be answered before examining the double taxation issue through the regime theory lens.

2. The History of the Double Taxation Regime

The current avoidance of double taxation regime has its origins in the 1928 League of Nations draft model bilateral income tax treaty and the documents preceding it (including the work of the International Chamber of Commerce (“ICC”).\textsuperscript{140} The central challenge, as understood clearly by the participants, was how to resolve the competing claims of the source and residence jurisdictions to tax the income from cross border transactions. On a theoretical level, two competing tax norms existed supporting the two sets of claims. Residence jurisdiction could be preferred on the grounds that: (1) it best reflects ability to pay (because the taxing state can “readily” base its taxation on the entirety of the taxpayer’s income and thus have an accurate sense of the taxpayer’s fiscal picture); (2) income “belongs” to people (residence), not places (source); (3) people are less mobile than activities; and (4) the source approach would put

\textsuperscript{139} The treaty provisions specify how states, through designated representatives, will evaluate and review problems under the treaty.

tremendous pressure on the definition of source. Alternatively, source jurisdiction could be preferred: (1) because of the benefits principle (the idea that the source country provides the infrastructure permitting the creation of the income); (2) because the source country may be aware of the income’s existence and hence better able to capture the tax; and (3) finally, because the source country can tax it.

As noted above, any state concerned that its own residents might face double taxation in their cross border activities could implement a foreign tax credit in its domestic legislation (relinquishing all or part of its right effectively to tax that income) and thereby substantially reduce the risk of such a burden. Alternatively, the residence country could exempt foreign source income from residence country taxation. The United States enacted a foreign tax credit in 1918, just five years after the introduction of the income tax system. Why would a country take this move unilaterally, which effectively allocates the first right to tax the foreign income of U.S. taxpayers to the source country? Three important reasons supported the decision: (1) it

141 See, e.g., Graetz & O’Hear, supra note __ at 1033-35.
142 See, e.g., id. at 1037-38 (Thomas S. Adams, a major U.S. tax advisor at the time of the framing of the international agreement on double taxation, expressed the view that source based taxation was inevitable because the revenue was “there” and it was unlikely that source countries could be convinced otherwise). This view has been repeated frequently. See, e.g., Stephen E. Shay, J. Clifton Fleming Jr., and Robert J. Peroni, “‘What’s Source Got to Do With it?’ Source Rules and U.S. International Taxation,” 56 Tax L. Rev. 81, 89 (2002) (noting claims that “force majeure” has driven source taxation as much as any moral argument).
143 See, e.g., Dagan, supra note __ at __.
144 If a country chooses to prevent double taxation of its residents through exemption of foreign source income, there is the possibility that the income will bear little or no income tax at all. For example, even if the source jurisdiction taxes the income item at a low or zero tax rate, the income would not be taxed in the residence country under the exemption system.
145 The decision to provide relief from double taxation through the credit rather than an exemption, seems to have been driven in part by the views of Thomas S. Adams, the proponent of the U.S. foreign tax credit in 1918. Adams saw residence countries as a valuable backstop in the taxation of income. If income escaped taxation in the source jurisdiction, it could still be taxed in the residence country under a credit approach. Graetz & O’Hear, supra note __ at 1038-1039 and note 73.
146 Edward Seligman, the influential U.S. economist who oversaw the League of Nations’ 1924 report on international taxation, viewed enactment of the foreign tax credit as “the United States . . . making a present of the revenue to other countries.” Edwin R.A. Seligman, DOUBLE TAXATION AND INTERNATIONAL FISCAL COOPERATION 133-34 n. 10 (1928) quoted in Graetz & O’Hear, supra note __ at 1046.
encouraged foreign investment (i.e. U.S. investment into Europe) which was desirable post-
World War I when U.S. private investment was viewed as essential to the rebuilding of Europe
and thus to Europe’s ability to pay its war debts;\textsuperscript{148} (2) it encouraged U.S. exports by increasing
European access to dollars;\textsuperscript{149} and (3) tax rates during and just after World War I were high
enough to turn double taxation from a nuisance into a serious issue.\textsuperscript{150}

After the U.S. implementation of a foreign tax credit several other countries followed
although not on the same scale (nor through the same mechanism), including Great Britain,
Belgium, Italy and France – all by 1928.\textsuperscript{151} Despite states’ ability to provide unilateral relief
through a credit (or an exemption), there was significant and widespread interest in developing a
treaty approach for resolving many of the issues not so readily subsumed under the foreign tax
credit. Important questions included: (1) how to determine the source of particular income; (2)
how to balance the concerns and desires of creditor and debtor nations (after World War I the
United States and the United Kingdom were major creditor nations and the continental European
countries were debtor nations)\textsuperscript{152} especially in rules governing interest and dividend taxation; (3)
what should be the limits on a state’s ability to tax the business enterprise of another state (the

\textsuperscript{147} U.S. tax advisor Thomas S. Adams expressed the view to the League of Nations
during the inter-nation treaty process that “[e]ach State should be eager, for selfish and economic reasons, to relieve
its own nationals and residents from that measure of double taxation which is due to its own legislation.” Graetz &
O’Hear, supra note __ at 1051 (quoting T.S. Adams).

\textsuperscript{148} See Graetz & O’Hear, supra note __ at 1049, 1052-53.

\textsuperscript{149} See Graetz & O’Hear, supra note __ at 1050 (quoting Mitchell Carroll in 1927: “The American
credit system is ideal for a wealthy nation that desires to encourage the expansion of its foreign trade, and is willing
to afford relief from double taxation to its own citizens or residents . . . . The United States says, in effect, to its
citizens – go abroad and trade” and we will provide the necessary tax relief.). In addition, where the foreign tax
credit facilitated cross-border investment, it encouraged economies of scale of production which would mean more
efficient production for the home market.

\textsuperscript{150} At lower rates of taxation, the effect of double taxation was more negligible. See Graetz &
O’Hear, supra note __ at 143-46;

\textsuperscript{151} Mitchell B. Carroll, Double Taxation Relief, Discussion of Conventions Drafted at the
International Conference of Experts, 1927 and Other Measures 1 (Dep't of Commerce Trade Information Bulletin
No. 523), 1927. at 2.

\textsuperscript{152} See, e.g., Graetz and O’Hear.
permanent establishment question); and (4) how to mesh the details and mechanics of two states’ domestic systems that begin with different tax bases and types of levies.153

The ICC, which was formed in 1920,154 directed early attention to double taxation. In 1920, the ICC adopted a resolution seeking “prompt agreement between the Governments of the Allied countries in order to prevent individuals or companies from being compelled to pay a tax on the same income in more than one country.”155 The ICC presented resolutions (not in treaty form) regarding double taxation at its 1923 Rome Conference, including features that remain with us today (classification of income by category for taxation, and allocation of business income among the source nations to which it corresponds).156 Ultimately, the resolutions were not adopted, largely due to British objections. The British opposition derived from strong British preference for residence-based taxation and their view that the resolutions granted too much taxation to the source countries.157

For a capital exporter, the British position is not surprising. Perhaps what is more interesting is that the United States, also a capital exporter in 1923, generally supported the substance of these resolutions. The U.S. position seems explained by: (1) Thomas S. Adams, the major (and influential) U.S. representative to the ICC, who had historically advocated for much source based taxation;158 (2) Adams’ interest in reaching international agreement on the details of double taxation (noted above) which required the support of debtor states;159 and (3) “the international balance of payments, which was overwhelmingly in the United States’ favor and

153 Several famous early treaties preceded the Models including: (1) 1924 Austria-Hungary Treaty; (2) 1922 Treaty of Rome (ratified by Austria and Italy); and (3) 1926 U.K.-Irish Free States Treaty. See Mitchell B. Carroll, Double Taxation Relief, Discussion of Conventions Drafted at the International Conference of Experts, 1927 and Other Measures 1 (Dep't of Commerce Trade Information Bulletin No. 523), 1927 at 2-4.
154 Cite- history volume.
155 Graetz & O’Hear supra note __ at 1066 (quoting the ICC).
156 Id. supra note __ at 1069-70.
157 Id., supra note __ at 1071-72.
158 Id. at 1072.
which permitted (perhaps even, in the interests of providing dollars for the purchase of U.S. exports and for the payment of U.S.-held debt, required) generosity in source rules to capital importers.”

As the ICC undertook its work in the early 1920s, the League of Nations issued its famous 1923 Report on Double Taxation by the “four economists” representing the United States, the United Kingdom, Italy and the Netherlands. The Report adopted a decidedly more pro-residence approach than the ICC by allocating interest and dividends to the residence (i.e. lender) country, and by failing to call for residence countries to provide a foreign tax credit. In fact, the structure of their recommended scheme mirrored their view that the double taxation question was not one of sharing and allocating taxing rights between residence and source countries. Rather, they understand the question as an exclusive choice – either residence or source jurisdiction for income. The 1923 Report divided taxes into those on global income (taxed only at residence) and all other taxes (taxed at residence or source depending on an “economic allegiance” principle that favored residence).

As its 1923 Report was being drafted, the League of Nations separately appointed a Committee of Technical Experts to translate the concerns over double taxation into more precise, practical and concrete suggestions. The Technical Experts issued their preliminary Report in 1925. In contrast to the 1923 League Report by the four economists, this 1925 Report reflected a

159 Id.
160 Id.
161 Id. at 1077-78. The work of the “four economists” was substantially led by Edwin R.A. Seligman, an American tax scholar of significant influence in the United States, but with a decidedly pro-residence stance in contrast to Adams. See, e.g., W.H. Coates, “League of Nations Report on Double Taxation Submitted to the Financial Committee By Professors Bruins, Einaudi, Seligman, and Sir Josiah Stamp,” 87 J. OF THE ROYAL STATISTICAL SOC. 99, 102 (1924) (citing the four economists long term hope that countries will come to the option of allocating the whole tax to residence (with an exemption system) as the most desirable resolution of the double taxation issue).

162 Id. at 1077. See Mitchell B. Carroll, Double Taxation Relief, Discussion of Conventions Drafted at the International Conference of Experts, 1927 and Other Measures 1 (Dep’t of Commerce Trade Information Bulletin No. 523), 1927 at 5-6.
stronger pro-source stance\textsuperscript{163} ("a majority of the drafters of the 1925 Report came from debtor nations"\textsuperscript{164}) and was endorsed by the ICC. The next step for the League was to draft an actual model bilateral treaty. In preparation for this assignment, the Technical Experts committee sought to broaden its drafting body.\textsuperscript{165} The United States came on board as one of the few creditor nations, motivated in part to: (1) protect its interests as such; (2) ensure a model that would integrate well with the U.S. system which did not use the more scheduler approach of many European countries; (3) encourage global uniform rules (to facilitate U.S. investment and trade); and (4) to encourage a multi-lateral treaty.\textsuperscript{166} The Technical Experts worked for several years, eventually producing the model treaty (actually three models reflecting different combinations of states with either unitary or mixed tax system) in 1928.\textsuperscript{167} The resulting allocation of income was described as one in which “neither the country of origin nor the country of residence makes a complete sacrifice in favor of the other.”\textsuperscript{168}

The next major contribution to the Double Taxation project came from the Organisation for European Economic Cooperation ("OEEC"), established in 1948 (predecessor to the OECD) as part of the U.S./Canadian aid distribution under the Marshall plan.\textsuperscript{169} From 1958-61 the

\begin{flushright}
\textsuperscript{163} For example, the 1925 Report granted certain categories of taxes to the residence or source country exclusively. Graetz & O’Hear \textit{supra} note __ at 1080.

\textsuperscript{164} See Mitchell B. Carroll, Double Taxation Relief, Discussion of Conventions Drafted at the International Conference of Experts, 1927 and Other Measures 1 (Dept of Commerce Trade Information Bulletin No. 523), 1927 at 5-6 (the seven original representatives came from Belgium, Czechoslovakia, France, Great Britain, Italy, Netherlands, and Switzerland). See, e.g., Graetz & O’Hear, \textit{supra} note __ at 1080.

\textsuperscript{165} These later representatives came from Argentina, Germany, Japan, Poland and Venezuela. T.S. Adams joined the group for the United States in 1927. See Carroll, \textit{supra} note __ at 5-6.

\textsuperscript{166} See, e.g., Graetz & O’Hear, \textit{supra} note __ at 1082.

\textsuperscript{167} Report Presented by the General Meeting of Government Experts on Double Taxation and Tax Evasion, League of Nations Doc. C:562 M.178 1928 II (1928); see also Graetz & O’Hear, \textit{supra} note __ at 1082-86;

\textsuperscript{168} Carroll, \textit{supra} note __ at 17.

\textsuperscript{169} See, Organisation for Economic Co-operation and Development: History of OECD: Organisation for European Economic Co-operation” at http://www.oecd.org/document/48/0,2340,en_2649_201185_1876912_1_1_1,00.html.
\end{flushright}
OEEC Fiscal Committee undertook the task of developing a model treaty because the business community perceived the existing (bilateral) double tax treaties to be inadequate. The OEEC Model was released in 1961 and reflected much consistency with the earlier League of Nations document.170 In 1961, the OECD (Organisation for Economic Cooperation and Development) superseded the OEEC and took over its Fiscal Committee and the treaty project. In 1963, the OECD issued its Model tax convention and commentary, which limited the source country’s ability to collect tax on investment income. Subsequently revised Models were released in 1977 and 2000 (with updates in the interim and thereafter).

Developing countries (i.e. capital importers favoring source jurisdiction) dissatisfied with the pro-residence orientation of the OECD model prompted the United Nations (which had replaced the League of Nations in 1945)171 to enter the realm of model tax treaty drafting. The U.N.’s Ad Hoc Group of Experts on Tax Treaties between Developed and Developing Countries tackled the problem of drafting a model to promote investment into the developing country while also ensuring good tax treatment in the residence country for the taxpayer.172 The Ad Hoc Group produced its first report in 1969, which reflected the OECD model as its starting point. In 1980, the first model was released.173 Although it favored source jurisdiction more than the OECD Model, it still reflected a compromise in that it granted neither source nor residence countries exclusive jurisdiction to tax. A more recent draft model was released in 2000.

The last major player to enter the drafting universe (although producing a draft before the United Nations) was the United States – with its 1976 Model. (Subsequent models included 1977, 1981, and 1996). Differences exist between the U.S. and OECD Models (reflecting particular concerns of the United States such as the treatment of citizens and the prevention of treaty shopping), yet once again the similarity between treaty models is dramatic reflecting a powerful commonality in terms of the perceived problem, the structure of its solution, and much

170  Cite - history and [].
3. **Regime theory application**

   Accepting the assertion at the beginning of Part II that there is indeed a regime governing the issue of double taxation, and that the regime comprises the above described activities, agreements, and documents, can we understand how it formed? Can we gain a better sense of when regimes will be successful in the international tax environment? To the extent that regime theory accurately describes this case study, it begins to define the contours of regimes in international tax. Of course additional steps (outlined in Part III) will be needed to develop confidence in our regime view of international tax. For example, we would need to review a wide array of cases, including those in which regimes failed to be established. Also, we would need to apply our regime theory model to cases currently in progress and attempt to predict the likely course of events. But this process begins here with the first case study.

   Following the two-step process outlined earlier, the first question is whether the double tax regime seems driven by power and distributional issues (neorealist tradition) or whether it seems to represent a case of market failure (neoliberal tradition).

   a. **Neorealist Regime Theory – Power and Distribution:** An initial reaction may be that any problem involving tax is inherently distributional because the rules are de facto allocating the rights to tax revenue to one country over another. However, to the extent that countries have approximately equal investment flows, then tax rules that are uniform (i.e. reciprocal) should ultimately provide the same net amount of tax revenue to each country. Of course at the very time that the first model treaties were being drafted the collected nations were not in fact equals but rather represented a set of creditor and a set of debtor nations. In such cases, the anticipated alignment of interests would be that debtor nations/capital importers would prefer a source-based rule structure and creditor nations/capital exporters would prefer a residence-based one (these preferences were not necessarily exclusive, but instead reflect which country they would favor in drafting the regime). This dynamic was mirrored in the clash between the United Kingdom and the drafters of the 1923 ICC Report “favoring” source based taxation. To the extent the double...
taxation issue is perceived as a distributional one, we might anticipate that the more powerful countries would prevail – by pushing for rules that favor their type of country. Given that creditor nations/capital exporters would typically be cast as the more “powerful” set of nations, we would expect the double taxation rules to favor residence jurisdictions. In contrast, the actual outcome suggests a degree of complexity not fully accounted for by a model based on power.

Recall that the United States and several other countries implemented a unilateral foreign tax credit (in whole or in part) – which favored the source jurisdiction as compared to the views espoused in the 1920s by the United Kingdom (advocating that most cross border income be subject to tax by the residence, not the source, country). Presumably these original unilateral foreign tax credit provisions were enacted based on the assessment (voiced by T. S. Adams\(^{174}\)) that a country benefited more by facilitating its own residents’ cross border business activity, than by collecting the taxes. Why? The answer reflects the temporal and economic dimensions of tax rules. On the temporal side, if no foreign tax credit is available this year, the residence country can obtain extra tax revenue, but in the future cross border activity will be discouraged with negative consequences for the revenue base and for the residence country economy. A country must measure its success by more than its current tax revenue;\(^{175}\) economic growth is also valued. If tax rules discourage taxpayers from engaging in otherwise desirable cross border transactions, there may be a significant drag on the residence economy and revenue. Thus, the divide between debtor and creditor nations in terms of their preferred tax rules was not as great as might have been anticipated. Although continental Europe was in debt, it had leverage (as explained by Adams) vis a vis creditor nations such as the United States because: (1) the United States wanted repayment (as opposed to offering debt forgiveness); (2) Europe was a viable and needed market for U.S. exports; (3) agreement on the details of double taxation would presumably facilitate cross border trade; and (4) the debtor nations were well represented in the

\(^{174}\) See supra note ___.

\(^{175}\) Or certainly more than its immediate revenue. If the residence country’s multinationals do business overseas, and the residence country allows a foreign tax credit, it may still get some tax revenue this year from those transactions, and it will continue (at least in theory) to see that flow. Also, the state will benefit from its industrial economy of scale (better goods at home too), and likely capture wages from producing companies
ICC and their support was needed for agreement upon any model treaty. The United States may have been powerful in economic terms, but for the issue area encompassed by double taxation, European partners were important actors as well.

What does this suggest about the distributional nature of the double taxation issue? Although there is a distributional aspect to the resolution of double taxation, particularly in the case of countries with unequal investment flows, it is not entirely clear that the distributional issues are as significant as the potential market failure problem described in the next section.\textsuperscript{176} That is, power does not seem to fully account for how agreement was reached and how the regime on double taxation was formed. This seems especially true where some countries (more powerful) took the unilateral step of enacting a foreign tax credit (which limited their revenue) before negotiating the remainder of the double taxation details with other states.\textsuperscript{177} Thus, although the model treaties offered by the United Nations, the OEEC, and the OECD tended to favor the residence country in its various source and allocation rules, the baseline idea of eliminating double taxation through a foreign tax credit (or exemption) favored the source jurisdiction.

Over time, other countries have entered the dialogue concerning the allocation of taxing rights between residence and source. Many of them are non-European countries that are in a distinctly different medium to long term economic posture vis a vis traditional capital exporters such as the United States, or even the capital importers of the 1920s (many of which were employees based in the residence country.

\textsuperscript{176} The double taxation issue can be envisioned as an effort to make the revenue pie bigger for countries by encouraging cross border transactions (and perhaps discouraging evasion of especially burdensome taxes).

\textsuperscript{177} However, one could argue that unilateral adoption of the foreign tax credit is neither evidence of power nor of market failure at the root of the double taxation agreement. If a certain practice is beneficial enough that it should be implemented, even if only unilaterally, then the fact that the regime was developed in two stages (first unilateral passage of domestic law foreign tax credits and then treaty work on a multilateral level) does not eliminate the power explanation. In this case though, the credit approach adopted unilaterally and promoted globally by the United States favored debtor/source nations and thus fails to resonate as a power move in regime formation.
countries recovering from World War I). Some of these new entrants to the debate balked at the OECD’s arguably pro-residence formulation of its model treaty provisions, a characterization that even the OECD accepted. As the countries had no direct mechanism for influencing the OECD’s model because of the “limited” nature of the OECD membership, they ultimately worked through the United Nations to produce a draft more sensitive to the position of developing, i.e. source, countries. This historical development leads to two observations

---

178 See, e.g., United Nations, Introduction to the U.N. Model Convention at viii (listing countries involved in the U.N.’s treaty drafting process, including Argentina, Brazil, Ghana, Pakistan, Sri Lanka, Republic of Korea, Mexico, Nigeria, and Venezuela).

179 At first it may seem contradictory to suggest that the behavior of the United States, or other countries that enacted a foreign tax credit, was “favorable” to source jurisdictions, while also noting that source/capital importing countries were displeased with the OECD model treaty. In fact both are true. Unilateral implementation of a foreign tax credit by a residence/capital exporter country does in fact seem to go against that country’s initial interest in tax revenue maximization. However, the OECD model treaty (as well as the other model treaties) encompasses more than the implementation of a credit (or exemption) method by the residence country. As we have seen, that can be accomplished unilaterally. What requires more negotiation are the complicated questions, including: Where are different items of income sourced? How are profits from a business allocated between or among countries? See supra text accompanying note ___. It is in these details that the OECD treaty was viewed as favoring the residence countries over the source countries.

180 See id. at vii (quoting the OECD: “the traditional tax conventions have not commended themselves to developing countries,” because “the essential fact remains that tax conventions which capital-exporting countries have found to be of value to improve trade and investment among themselves and which might contribute in like ways to closer economic relations between developing and capital-exporting countries are not making sufficient contributions to that end . . . Existing treaties between industrialized countries sometimes require the country of residence to give up revenue. More often, however, it is the country of source which gives up revenue. Such a pattern may not be equally appropriate in treaties between developing and industrialized countries because the income flows are largely from developing to industrialized countries and the revenue sacrifice would be one sided.”)

181 Developing countries could, in theory, have started afresh on the issue of double taxation, rather than rely on the existing double taxation treaty framework. However, such an approach would have entailed a number of costs, including the developing of the alternative structure, and the loss of familiarity, certainty and predictability of the longstanding framework (as their new approach would have a shorter history behind it). Tax treaties appeal to developing countries for a number of reasons, chief among them the comfort that they provide to the new investor. This comfort can be both concrete and intangible. On the concrete side, treaties facilitate the intersection of two countries’ tax system and provide a framework for resolving conflict. On the intangible side, treaties can signal to investors that a country is part of the “international” system and one can be comfortable pursuing business and investments there. Many emerging market countries believe that a treaty is an important signal to potential investors about the status and reliability of the nation – i.e. you need to have a treaty to be perceived as in the plausible and viable investment destination. See, e.g., Lee Sheppard, “Berman, Part II: Departing U.S. Treasury Staffer Discusses Treaties,” 15 TAX NOTES INT’L 949, 950 (Sept. 22, 1997) (quoting Deputy International Tax Counsel Daniel Berman: “It is often intimidating for U.S.-based companies to do business in a foreign country knowing that if they have a problem with the tax system, they’re going to be at the mercy of the
regarding the neorealist understanding of regime formation.

First, consistent with the neorealist view of what functions a regime can perform, the double taxation regime and its reliance on model treaties drafted by different international bodies, allowed the weaker states (post-World War II developing countries) to voice their views on treaty structure. The implication is that participation inside the regime with the potential for some influence is preferential to sitting outside the regime entirely. Along these lines, it is interesting to note that the United Nations used the OECD model as its starting point (i.e. not only were these developing countries working inside the double taxation regime, they were really working inside by accepting the pre-existing structure for resolving the issues).

Second, the double taxation regime exists separate from and independent of the continued participation of many developing and emerging market countries. In that sense, the regime at its core is not about a major power/distributional struggle in resolving double taxation. That said, to the extent the basic model treaty provisions favor residence countries (although that has minimal effect in the case of equals) it does disadvantage developing countries because it layers the tax treatment on top of the pre-existing economic and investment imbalance between the two countries. Why do the developing countries continue, for the most part participate in (and in fact often clamor for) bilateral treaties even on these general terms? As described below, the

---

182 But also not inconsistent with the neoliberalist perspective considered in the next section. See supra note [180].

183 See supra note [180]. See United Nations, Introduction to Model Treaty at x (the United Nations Ad Hoc “Group of Experts . . . decided to used the OECD Model Convention as its main reference text in order to take advantage of the accumulated technical expertise embodied in that Convention and the Commentary thereon, and also for reasons of practical convenience stemming from the fact that the Convention was being used by OECD member counties in the negotiation of tax treaties not only with each other but also with developing countries.”). See supra note [180]. The OECD and the U.N. Model treaties also overlapped in personnel as well. Stanley Surrey, U.S. Assistant Secretary for Tax Policy (1961-69), and one of the “dominant player[s]” in the 1960s in tax policy, “was also a major player in international organizations like the OECD (which issued its model tax treaty in 1963 . . . heavily influenced by Surrey) and the UN (Surrey was the major force behind the UN 1980 model tax treaty).” Reuven Avi- Yonah, “All of a Piece Throughout: The Four Ages of U.S. International Taxation,” 25 VA. TAX REV. 313, 327 (2005).

184 See Treasury stats on pending treaty negotiations and treaty requests. But see the history of the
treaties provide benefits that the market has failed to – coordination of these tax rules between nations. Thus, without rejecting the existence of some power and distributional elements on the formation of the double taxation regime, the case study does not seem adequately explained by reference to these factors. The next section considers how a neoliberal view of regime formation illuminates the activities described above.

b. Neoliberalist Regime Theory – Market failure explanations:

i. Bargaining Game Model

In order to assess the applicability of the neoliberal model, this part begins by trying to identify the type of bargaining game involved. Recall the idea is that the states are, in the absence of the regime, not engaged in “optimal” behavior. Without cooperation the parties’ choices would not be pareto efficient because there would be alternative outcomes (decisions/choices) that would make at least one better off, without making the other worse off. However, moving to that pareto efficient outcome requires some measure of cooperation. The expectation is that in some cases formation of a regime can provide the needed cooperation. The type of “game\textsuperscript{186}” being played impacts that likelihood that regime formation will be possible.

Four basic types of games were outlined in Part I:

(1) Suasion: The central idea in a suasion game is that the two parties have differing incentives and one party has to be persuaded to “play.” In the “cooperation version” of the suasion game, the dominant strategy for party A is to cooperate –thus regardless of whether party B decides to cooperate or defect, A will cooperate. But the best outcome for B is if A cooperates and B defects, and this can be achieved automatically because in fact A’s dominant strategy is to defect. Essentially, A (who wants B to cooperate) has two choices: (1) threaten to act irrationally (i.e. threaten to defect even though that is against A’s best interest), or (2) offer something extra to B if B cooperates on this matter. It can be difficult, though not impossible, for A to persuade B that is will act against its

---

\textsuperscript{186} As the earlier examples reviewed, the different types of games are meant to capture the different types of incentives and or strategies that countries may have in negotiating over cooperation. See supra text accompanying note \textsuperscript{__}.
own interests. However, it is possible to offer B something else for cooperation. Thus, a
suasion game on one issue may be linked with other suasion games between the parties in
some type of trade.

(2) PD/collaboration: characterized by the prisoner scenario\textsuperscript{187} that lends its name to this
game – the best mutual outcome (i.e. the pareto optimal outcome) for the parties (two
prisoners who committed crimes together) is joint cooperation (silence to the police), in
which case they get a light sentence. But the worst outcome is to cooperate (be silent)
while the other party defects (confesses to police and receives no sentence) because the
cooperating prisoner receives a heavy sentence. If however, both defect (confess), they
do not go free and instead receive a medium sentence. In absence of a mechanism to
ensure the other party will cooperate, the best strategy is to defect (confess) – leading to
mutual defection (and a medium sentence) although that is not a pareto optimal outcome.

(3) Coordination: characterized by the air traffic controller example – the best outcome for
the parties is cooperation, but unlike the PD scenario, here there is no risk of defection
once some agreement has been reached. (For example, once a language for air traffic
control has been selected, no one has an incentive to defect. If the selected language is
Italian, a country does not benefit by refusing to speak Italian or by having its
controller/pilots use another language. That behavior would only court disaster.
Correspondingly, the parties do not fear defection on the other side). Thus, in a
coordination context, the agreement is “self-enforcing.”\textsuperscript{188}

(4) Assurance: Typified by the strategies of two stag hunters -- both hunters prefer to
cooperate and hunt the stag. However, mutual defection is also an equilibrium because if
one hunter defects, then the other should as well (both obtain rabbits). The worst
outcome is if one hunter defects (goes after the rabbit) and the other hunter continues
hunting for the stag alone (and thus catches no stag).\textsuperscript{189}

\textsuperscript{187} See supra note ___.
\textsuperscript{188} See supra note ___.

47
Application to Double Taxation – the General Case: As an initial matter, if we take cooperation to mean that a country unilaterally provides a foreign tax credit or exemption of foreign source income\(^{190}\) and defection to mean a country refuses such a credit or exemption, then the parties arguably prefer joint cooperation as the best outcome.\(^{191}\) Imagine two countries, A and B. In the absence of any coordinated strategy, how will A and B each respond to the existence of double taxation? Assuming both A and B view double taxation as undesirable (because each wants its residents to engage in cross border business to stimulate its economy and enhance the economic well being of the country and its residents),\(^{192}\) then they would relieve double taxation for their residents regardless of the other country’s decision. Thus, even if the country B does not relieve the double taxation of its residents, country A would provide the credit/exemption for its own residents. Why? Because A’s loss of tax revenue (from the credit/exemption) is offset by the economic benefits of having A’s own residents pursue cross border business opportunities. Yes, country B may, at least in the short run, have some extra tax revenue, but B’s residents will likely reduce their cross border activities which negatively impacts B’s tax revenues and B’s broader economic picture. Thus, cooperation, even in the face of defection, is the preferred strategy for A and for B.\(^{193}\)

Moreover, cooperation (i.e. the use of a credit or exemption) by both A and B is probably

\(^{189}\) See supra note __.

\(^{190}\) As discussed earlier, both the foreign tax credit and the exemption method alleviate double taxation by advantaging source-based taxation. It is also possible to reduce double taxation by granting more taxing rights to the residence country over source. However, given that a primary motivation behind the double taxation movement in the 1920s was a concern that one’s own residents be able to pursue cross border business activities without undue (or perhaps crippling) tax burdens, countries willing to unilaterally try to resolve the double taxation problem could only do so by relinquishing their own rights to tax – granting the foreign tax credit or providing an exemption. Thus, the “cooperative” strategy at this stage implies granting a foreign tax credit or exemption– which also relinquishes residence country taxation in favor of the source country. Recall, that the main difference between the foreign tax credit and the exemption method

\(^{191}\) That is, joint cooperation (both countries granting foreign tax credits) would significantly reduce double taxation.

\(^{192}\) See supra text accompanying note __.
their preferred outcome. That is, of the four possible results (both A and B cooperate, both A and B defect, A cooperates and B defects, A defects and B cooperates), they would prefer the outcome where they both cooperate. Why? For several reasons: (1) Taxpayers in country A may, under some circumstances, be treated as residents of country B (if, for example, the country A taxpayer had a subsidiary in country B). In such cases, A would prefer that country B relieve double taxation for its residents because A’s grant of a foreign tax credit or exemption would not help the subsidiary in country B earning income in country C; (2) Mirroring assumptions from the trade arena, country A may adopt the view that it benefits when the global economy is more open and investments and activities can freely move around the world. If country B does not relieve its residents’ double taxation, that may hinder this economic mobility, thus leading B to desire A’s cooperation here in the provision of unilateral relief for residents.\textsuperscript{194}

In terms of possible outcomes, joint defection is the worst. If neither A nor B relieve their residents double taxation then each country’s own residents will not pursue valued global commercial activities (triggering corresponding national economic and revenue effects), and the global economy overall may suffer. Implicit in the above analysis is the conclusion that an outcome in which one cooperates and the other defects is less desirable, but not as undesirable as mutual defection.

Thus, the four possible outcomes can be ranked in terms of A’s likely preference (best to worst):\textsuperscript{195} (1) joint cooperation; (2) A cooperates and B defects; (3) A defects and B cooperates; and (4) both A and B defect. An important assumption made here, that will be relaxed later, is that A and B both share the principle that double taxation is harmful and should be avoided.

\textsuperscript{193} Preferred strategy here indicates that it is the action the party would likely choose in the absence of any coordination.

\textsuperscript{194} Of course country A could provide some limited relief to the double taxation faced by country B residents in those cases where the B residents are earning income in country A by exempting them from source country (i.e. country A tax). This outcome requires A to highly value the global free flow of investments given that the investors are not their own residents, or to strongly desire the investment into country A. However, even if country A would consider offering this relief, it presumably would prefer that country B resolve the double taxation of B’s own residents.

\textsuperscript{195} B’s preferences would be the same, except the labeling of A and B in the listings would be reversed.
They then adopt the norm that residence countries should yield to source (this adoption may be either a function of unilateral practicality or of their views of legitimate taxing jurisdiction). Finally, the rules they use to implement their norms and principles here (the foreign tax credit and exemption system) are sufficiently compatible without further accommodation. The assumption that A and B share a similar vision of the principles, norms, and rules is realistic where both A and B are developed countries. However, if A is a developed country and B is a developing country, then the assumption must be re-evaluated. The next section examines the “game” where A is a developed country and B is a developing country.

Based on the above analysis and rank ordering of outcomes for A and B (as two developed countries), what type of bargaining game characterizes the double taxation scenario? The suasion game is eliminated because this is not a case where there is a single equilibrium which makes one party happy but not the other. The PD (prisoner’s dilemma) is also eliminated because the outcome that results in this double taxation case without any mutual agreement is the pareto optimal outcome. A coordination game model fails to describe the double taxation case because, under the core assumptions in the problem, the choice is either to cooperate (institute a unilateral foreign tax credit or exemption system) or defect (offer your residents no relief from double taxation). Both A and B can arrive at rules (an implementation) that will work without negotiation between them (they can choose either the foreign tax credit or exemption system, and effective cooperation will follow). In contrast, the air traffic controllers example revealed agreement on the need for a single language but differences in the implementation (which language). Finally, an assurance game characterization also fails because in such a game (e.g., stag hunt) both parties favor the same outcome (cooperation – getting the stag) but also favor joint defection (getting two rabbits) over a cooperation/defection pattern (getting one rabbit). In contrast with double taxation, country A (or B) would prefer to “cooperate” even if the other state defects, as evidenced by the unilateral foreign tax credit

196 If a country strongly holds the principle that international double taxation of its residents should be avoided, the only unilateral option open is for that country to offer its residents a foreign tax credit or exemption.

197 See infra text accompanying note __.
examples.

At this point, it looks as if none of the categories applies, and that the double taxation problem is not a game theory model for which the parties’ need help in achieving desirable outcomes. If that is true, why would countries have pursued an elaborate process of developing and negotiating model and actual income tax treaties to relieve double taxation? It appears that both country A and country B would unilaterally come to the same conclusion and adopt double taxation relief without reliance on any “regime” to facilitate the process.

The answer is that this view misses a key point: there are in fact two levels of “cooperation.” The first level of cooperation concerns the decision to offer double taxation relief for your residents – an outcome that seems to flow from unilateral decision making. The second level of cooperation, however, concerns optimization of the double taxation relief system.

Although the paper asserted that no coordination of the rules was required in relieving double taxation, and that country A could adopt the credit and country B could adopt the exemption method – in fact that is an overstatement. In theory A and B can adopt mechanisms unilaterally that should provide adequate relief from double taxation. In reality, however, the likely intersections of their tax rules (such as tax definitions and source rules) require some coordination to provide maximum relief from double taxation. To the extent differences remain on these elements, some double taxation will exist, even if, for example, both countries grant foreign tax credits. If we take the provision of a foreign tax credit/exemption as a baseline, then the game is really about the implementation details and the coordination necessary to ensure full relief. Countries may be willing to take these details seriously and devote resources to their resolution where the double taxation situations are essentially iterative games as many taxpayers from the two countries keep facing the intersection of the countries’ rules.

Having established that the treaty regime is not really about the basic implementation of

---

198 See supra note __.

199 Under such facts, one could argue that a fifth choice from the IR literature applies: a harmony game. In a harmony game, both sides unilaterally prefer, and thus cause, the cooperative outcome without the need for a regime. This game was not discussed earlier in the paper because it provides little guidance on regime creation. Moreover, it clearly was not relevant here as the regimes were created to develop agreement on the
double taxation relief (that would be achieved unilaterally) but rather about the more sophisticated question of ensuring complete relief in the face of conflicting details, we must now determine what kind of game is in play. The best analogy would appear to be the air traffic controller example, a coordination game. Both country A and country B would prefer “cooperation,” here understood to be the use of shared definitions and detail rules. Some mechanism is necessary to reach agreement on these rules (similar to the selection of the air traffic control language)\(^{200}\) and once they are set, the parties do not have a very strong incentive to defect from the definitions.

One notable distinction separates the double taxation case from the classic coordination game of the air traffic controller, although ultimately this distinction should not be enough to render the coordination classification invalid. Unlike the air traffic controller example,\(^{201}\) it is possible that countries might perceive a defection from certain double tax rules and definitions to be to their advantage if this defection increases their tax base, while the other country continues to follow definitions which limit its taxing power. In truth, such defections are not in the country’s medium to long term interests given two connected facts: (1) double taxation is an iterative game, and (2) defections will generally be transparent. If country A defects, this defection will be transparent because it will immediately affect country B residents who will be very unhappy about their tax treatment. They will face unexpected double taxation because, for example, country A is taxing an item of income that had been deemed sourced in country B under the treaty. If country A is taxing an item of income that was considered allocated to country B, country B will typically still tax that item (as it is entitled to do). The resulting double taxation will trigger two responses. First, the residents of country B will hesitate to invest in country A because the existence of the treaty is not protection against double taxation.

---

\(^{200}\) Just as having a country’s native language selected for the uniform standard can be desirable, so too having tax rules and definitions that resonate more clearly with a country’s domestic tax rules or that seem to grant the country some additional potential revenue can be desirable. Countries will not naturally and unilaterally reach the same conclusions on these issues, and thus the regime is necessary to facilitate the reaching of an agreement on the details (language or tax rules).
Second, country B may seek to formally end the treaty or may engage in the kinds of informal defections that country A pursued. Either or both of these outcomes is undesirable for country A, thus A is unlikely to defect from the agreement.  

How difficult will it be to reach this agreement, from which neither A nor B will likely defect? Assuming, as we have, that both A and B are developed countries with comparable investment flows, then mutual cooperation should achieve improved trade and investment activity. If flows are equal between the states, then regardless of whether the decision is made to favor the source or the residence country in the allocation of taxing rights, the net revenue effect to the states should be the same. Thus, the distributive effects of the rules would be relatively minor, along the lines described above, and agreement quite probable.

Application involving a developing country: The observation regarding the ease of reaching agreement, brings us back to the central assumption in the analysis of country A and country B – that they were both developed countries. How does the analysis change if country A is a developed country and country B is a developing country?

Returning to the threshold

---

201 A country’s decision to defect from the selection of Italian as the universal air traffic controller language will likely result in the crash of a domestic plane in another country, or a foreign plane in defecting country. Neither result is desirable, even in the short run, and hence defection is very unlikely.

202 Although defection is unlikely, that does not mean that countries would not complain or lobby for a change in the rules (i.e. language selected).

203 This is not to suggest that countries of comparable investment flows will have no conflict in negotiating their double taxation treaties. Some conflict is quite likely – both because flows and economic circumstances will never be identical, and because countries may have reasons beyond direct economic impact for preferring some rules over others. That said, the more similar the economic status of the two countries the fewer distributive effects the rules will typically generate.

204 It is also possible that country A and country B are both developing countries. The text does not elaborate on this scenario because it is less likely to lead to any negotiations over a treaty, and where it does, the structure of the dynamics more closely mirrors that of two developed countries. If both A and B are developing, then they primarily seek investment, which neither of them has to offer. Neither A nor B is eager for its own residents to invest abroad instead of at home, and so at a unilateral level they will be inclined to offer no unilateral relief. A and B are not seriously concerned about whether the other offers relief to its own residents because those residents have nothing to invest. For example, country B does not care if country A offers relief; country A residents were never going to be serious investors in country B (they have little to invest) and so the added drag on investment from double taxation is not significant. At the margin, country B might prefer that country A offer relief, just in case a country A resident were considering an investment in B. However, that preference is relatively weak and not likely to prompt B to pursue treaty negotiations. Plus, in negotiation, A would demand that B offer relief as
cooperation decision of whether to provide double taxation relief or not, A’s ranking should remain the same. The dominant force behind A’s desire to relieve double taxation for its residents was the view that such relief improved the economic and revenue picture for country A.\textsuperscript{205} Country B, however, may have a quite different perspective if: (1) it has few residents that are ready and able to pursue cross border investment opportunities (or even domestic ones), and (2) greatly desires investment into its country. Under these assumptions, B’s rankings (from best to worst) on the basic cooperation question (relief or no relief) would be: (1) A cooperates and B defects; (2) both A and B cooperate; (3) both A and B defect; and (4) A defects and B cooperates. This ranking reflects B’s priority of getting (and keeping) investment in its own country. If country A cooperates and B defects, then the residents of country A can invest in country B without fear of double taxation (“good” for country B). At the same time the few residents of country B with investment capacity will, at least to some degree, prefer investment in country B to avoid the double taxation that will occur with investment outside country B (“good” according to B’s stated goal of increasing investment in B). If both A and B cooperate, then country A residents will invest in B (“good” for B) but country B’s few resident investors will invest (to some degree) outside country B (“bad” according to B’s identified goals). A still less desirable result follows from joint defection because if A offers no relief from double taxation, then country A residents will not invest in B (“bad” for B), although at least B’s few resident investors will be inclined to invest in B (“good” for B). Least desirable is the outcome in which A defects and thus its residents will not invest in B (“bad” for B), and B cooperates and

\footnote{These expectations seem reflected in the United States’ unilateral decision to implement the foreign tax credit in 1918. See \textit{supra} note \textsuperscript{205}.}
its residents invest outside B (“bad” for B).

Two observations emerge from this effort to identify the preferences where A and B are developed and developing: (1) the likely outcome, in the absence of agreement between A and B is: country A cooperates and country B defects, and (2) B does not just prefer an outcome in which A cooperates and B defects. It actually matters to B how A achieves relief of double taxation. If A selects the exemption method, then B has an opportunity, should it choose, to levy little or no source country tax in order to encourage country A residents to invest in B. Conversely, if A selects the credit (at least in the absence of tax sparing), B gains little or nothing by imposing a rate of tax lower than country A because A will first impose its tax on country A residents’ income and then permit a credit for any country B taxes.

Arguably the resulting outcome, A cooperates and B does not, is a pareto optimal result. No one’s position can be improved without harming the other: the outcome reflects B’s first choice, and to the extent A would prefer that they both cooperate, a move in that direction would make A better off, but would make B worse off. In terms of a “game” type, the developed/developing country pairing looks like a suasion game where one party wants to defect and the other has a dominant strategy to cooperate. Thus, we might anticipate that A would have to entice B into “cooperation.”

In truth, however, just as with the developed/developed country pairing, there are two levels of decision making. The preliminary decision concerns whether to offer unilateral relief.

---

206 The foreign tax credit provides a benefit to the residence country taxpayer only when the source jurisdiction actually levies an income tax. If the source jurisdiction imposes little or no tax (to attract investment), the residence country taxpayers would still face residence country taxation on the foreign source income (either now, if operating in branch form, or later, if operating in the source country through a subsidiary). Thus, the source jurisdiction is limited in its ability to use a low tax rate to attract investment. To counter this effect, some residence countries have implemented tax sparing provisions to assist source countries that are also developing countries. Under tax sparing, if a resident earns income in the developing source country and that country levies no income tax, the residence country will still provide a credit for a “phantom” tax deemed paid to the source country. Thus, the residence country investor benefits from investing in the no tax source country, and presumably will be drawn to invest there as a result.

207 Country B can benefit from implementing a zero or low tax rate in two ways. First, if country A residents invest in country B through a country B subsidiary, then country A tax may be deferred until the subsidiary pays a dividend (assuming no anti-deferral rules apply). Second, country A must be aware of the income earned in country B in order to effectively tax it. That is, if country A residents are willing to evade tax otherwise due to country A on their foreign source income, then a zero tax source country would be appealing.
On that, the developed/developing country pairing produces equilibrium. However, in reality both A and B need more to fully achieve their goals. In order to ensure that there is no double taxation, A wants to work out with B the details of the taxation of country A residents earning income in B. The step requires a regime in which they can reach agreement, the treaty mechanism. But again this raises the obvious question, Why would B participate? Recall that B’s driving motivation was not relief of double taxation, but rather the encouragement of investment into country B. A small amount of investment will come from country B resident investors who feel forced to invest in B because they face double taxation on foreign source income (because B offers no credit or exemption). Country A investors, who bring the bulk of investment into B, will not come if they face double taxation or uncertainty. Given B’s goal to increase investment, B will strive to reduce double taxation and uncertainty.

As to double taxation there are two ways to ensure country A residents will not suffer: (1) negotiate a treaty and iron out the details with country A, or (2) not tax country A residents (i.e. no source country tax). How should B decide? A treaty would allow B to actually collect some tax without driving A residents away. Alternatively, a zero tax could serve as strong attraction for investment (at the cost of revenue) – but typically it will require that the country A resident invest in B through a country B subsidiary. This tradeoff suggests that country B might not care about pursuing a treaty, at least if it is willing to forgo taxing residents of A. However, if it wishes to tax residents of country A or if it wants to address the second reason that A residents will not invest (uncertainty), then a treaty may be desirable to country B. How does the treaty resolve uncertainty? It does so explicitly and implicitly. Treaties explicitly ameliorate uncertainty by addressing questions of potential ambiguity or rule overlap, and by providing a framework through which to resolve conflict with country B using the help of country A (the competent authority mechanism). Treaties can also implicitly reduce uncertainty by signaling

---

208 Obviously, investment in B by country A residents is not subject to an on/off switch. Even in the face of uncertainty and/or double taxation, some country A residents may invest in B. What can be said is that uncertainty and double taxation will seriously dampen enthusiasm for investment in B, and may also limit the types of investments that are considered attractive.

209 See supra text accompanying note ___ (explaining the circumstances in which a country A resident can defer A tax on its foreign source income, and thus benefit in the interim from a low or zero tax rate at source).
that country B, a developing country, is playing according to well known and established rules (as spelled out in the treaties) and has already been able to negotiate and work with the residence country (as demonstrated by the treaty negotiations), and therefore is an viable investment destination.

Thus, for at least a significant subset of developing countries, their real preference is in fact cooperation (i.e. relief of double taxation and full resolution of the details) by both A and B. Although their desire for investment might initially lead them to favor an outcome of cooperation by A and defection by B, a fuller consideration of their investment goals will lead them to “embrace” the principle of eliminating double taxation. Moreover, they will also share the norm that the residence country should yield to source, as they are generally capital importers and hence their dominant role is as a source country. The details, however, may prove more troublesome, as described below, because many choices will not be a wash because the investments flows between the two countries are not equal. These developing countries will be willing to move forward with the negotiation of a treaty – although in that process, their status as developing countries will influence their views on the rules and the ease with which agreement will be reached.

The remaining subset of developing countries, those for example that are primarily interested in zero taxation and who believe that they are sufficiently well known to investors to be attractive jurisdictions, may stop at the first decision making level (unilateral determinations on double taxation relief). Country A will provide relief (credit/exemption) and country B either will provide no relief, or will provide it but not seek to coordinate with country A.

Returning to the subset of developing countries that seek to negotiate a treaty, how and why exactly will the selection of rules be difficult? In a treaty context, a given country is both a residence country and a source country. If the flows between the two negotiating countries are equal then a treaty that favors the residence country or the source country does not in reality favor either of the two countries. However, if A is a developed country and B is a developing country, the choice of rules\textsuperscript{210} to accomplish the elimination of double taxation will not be

\textsuperscript{210} That is, the specific method of “cooperation..”
revenue neutral (as compared to a pairing of developed countries). In terms of real investment flows, the developed country will be predominantly the capital exporting country, i.e. the residence country, and the developing country will be predominantly the capital importing country, i.e. the source country. As an example, recall that in the 1920s, debates over the various formulations of the allocation of taxing rights to source or residence countries were significant where the countries perceived themselves as being creditor or debtor nations. The creditor nations would be mostly affected by residence country rules and the debtor nations would be mostly affected by the source country rules.

Under these circumstances, the coordination game has a strong distributional component (in contrast to the coordination game example of air traffic controllers). Although B has made the decision to provide and coordinate double taxation relief (unlike the other subset of developing countries), it may be hesitant to cooperate in the second level of the game if coordination is understood to mean adjustments that favor developed countries. Country B would need to determine whether the loss of revenue from acceding to those adjustments would be offset by increases in valued economic investment and activity. The reaction of many emerging and developing countries to the possibility of a tax treaty with the United States suggests that on balance they view having the treaty as desirable for the reasons identified earlier: (1) the business community is comforted by the existence of a treaty, (2) treaties signal a country is part of the global economic and business mainstream, (3) treaties provide a procedural bridge (e.g. for dispute resolution), and (4) treaties provide predictability by answering questions on the intersection of the two countries systems. If here the regime could develop more context specific rules, it could facilitate agreement between developing and developed countries on many of the detailed matters necessary to eliminate double taxation. The work of the United Nations, falling on the heels of the OECD model (which gained broad acceptance), can be

---

211 This second level involves coordination of the definitions and adjustments to the allocation of taxing rights between the source and residence countries.

212 This point actually reflects the neorealist idea that regimes could perform certain functions including allowing differential treatment for “weaker” states.

213 See supra text accompanying notes ____.
understood as playing this role to the extent it provided an alternative yet sufficiently familiar structure for coordination involving countries with unequal flows.

Thus, although complex factors dominate the double taxation case, the coordination game seems a relatively strong description of the nature of the interactions. As such, we might expect better success at regime formation than in the case of the suasion games or PD games.\textsuperscript{214} Although extensive monitoring may not be necessary, defection cannot be as readily ruled out as in the true classic example of coordination games – such as the air traffic controller language selection.\textsuperscript{215}

ii. Regime Topic

After determining the nature of the bargaining game involved, the next step (under a neoliberalist approach) is to consider the impact of the topic itself on the likelihood of regime formation. Conflicts are disagreements over one of the following: (1) means, (2) values, (3) relatively assessed goods, and (4) absolutely assessed goods. As a general matter there seems little disagreement that double taxation is a problem (although the measures taken to eliminate it may reflect differing views on the severity of the problem relative to other concerns). This conclusion is reflected in the observation that in a developed/developed pairing, both A and B will unilaterally “cooperate” and provide relief. In a developed/developing pairing, those developing countries anxious to pursue a treaty have made the calculus that overall reduction of double taxation is desirable (even if they are not so driven to provide relief for their own residents.). Thus, there seems to be no serious debate over values.

Does the double taxation issue involve a conflict over means? Again countries seem to agree that one jurisdiction (residence or source) must surrender taxing rights in some

\textsuperscript{214} The assurance game, which was rejected as a characterization would expect easier regime formation. See supra text accompanying note __.

\textsuperscript{215} Certainly examples of defection can be identified, such as U.S. treaty overrides. However, such defections are highly public and do not raise monitoring issues. In addition, U.S. law seeks to minimize the likelihood of treaty overrides, at least explicit ones. See I.R.C. § 7852. See also, S. Rep. No. 100-445 (100\textsuperscript{th} Cong., 2d Sess. 1988) (“neither the treaty nor the [U.S. tax] law shall have preferential status by reason of its being a treaty or a law... [t]he committee does not intend this codification to alter the initial presumption of harmony between, for example, earlier treaties and later statutes.”). Ultimately, defections of this sort can be viewed a unilateral renegotiations of the rules that will risk triggering retaliation. The decision of the treaty partner on the retaliation question requires an independent analysis of risks and benefits.
circumstances. Where the tensions arise is over which jurisdiction that will be and to what extent. That debate is not founded on a disagreement about whether favoring source or favoring residence is actually more effective at achieving the goal of eliminating double taxation (a question of means), but rather on an allocation issue of who may have to bear any burden for achieving this goal. Unless the flows between the countries are equal, favor source or residence will impose some burden on one of the countries. Thus, the conflict concerns goods, here, the tax revenue.

This leaves the question of whether the conflict is over relatively assessed or absolutely assessed goods. Given the distinction turns on whether a country’s enjoyment of the goods correlates with how much of the goods the other country has, the application in the double tax context is not completely self-evident. First, tax revenues are money and money of course is power, as well as the ability to buy goods of any type, whether relatively or absolutely assessed. This observation provides little initial guidance as to the classification of the conflict. Second, where the countries involved have relatively equal investment flows, an agreement of any type should improve those flows for both parties, which is a benefit. Whether the specifics of the agreement favor residence or source should not have a significant distributional effect on the parties. Described in this way the conflict seems less about goods and more about the means (i.e. the details as to which choice of rules is more effective), as noted above. In that case, the conflict is of the type most conducive to regime formation.

However, to the extent the parties are not equals, for example A is a developed country and B is a developing country, tax dollars are at stake and we return to the money as power characterization. Tax revenue is liquid and could readily become any asset the state sought. Thus, the double taxation issue could be viewed as a conflict over relatively assessed goods because countries value revenue in part by reference to how it compares to that of other

---

216 Although there was discussion on a more theoretical level about the relative merits of source and residence countries’ claims to tax revenue, the core double taxation debate seemed more of an allocation battle between capital importers and capital exporters, and later between developed and developing countries. See supra text accompanying notes ___.

217 See infra text accompanying note ___.

60
countries. 219 If this assessment is accurate then the likelihood of regime formation for a developed and a developing country is less than for two developed countries, where the conflict could be characterized as a conflict over means. 220

Is it accurate, though, to view a dispute over revenue as a conflict over relatively assessed goods simply because a country could purchase such goods with its tax dollars? Food, which is often considered an absolutely assessed good, could in theory (and in reality) be used to buy guns, which are typically considered a relatively assessed good. Does this “substitution” argument undermine any distinction between conflicts over relatively and absolutely assessed goods? Despite the blurring of the distinction between the two categories of goods because of their exchangeability in the market, the distinction between relatively and absolutely assessed goods may still be a valuable measure of the likelihood of agreement on an issue. In the actual negotiation process the nature of the good literally “on the table” might influence the negotiations (perhaps because of its salience and proximity). Thus, even though food can be used to buy guns, a dispute over foods may nonetheless be less sensitive in a concrete negotiation than a dispute over weapons.

Where does this leave the categorization of the double taxation conflict, and how much does it matter? If the double taxation conflict is viewed as conflict over absolutely assessed goods, then the likelihood of reaching agreement is better than if we believe it more closely mirrors a conflict over absolutely assessed goods. If it constitutes a conflict over means, then as noted, agreement should be easier to achieve. Our labeling here does not make negotiations over revenue easier or harder, but a more accurate understanding of the conflict may improve understanding of the negotiation process. The history of bilateral tax treaties supports the idea that for countries of comparable economic situation (e.g., two developed countries) the double taxation issue may most closely resemble a conflict over means, which are the most readily

218 See supra text accompanying note ___.
219 See, e.g., different statistics measuring and comparing countries on a number of economic and dollar flow dimensions.
220 See supra text accompanying notes ___ (outlining the ranking of conflict types in order of the
formed agreements. When the United Nations and the developing world looked at the then existing model treaty (from the OECD) in the 1960s, they concluded that the terms were not easy for a developing country to accept, although they had worked adequately for the developed world. From the perspective of the developing world this was no “means” dispute, serious tax revenue was at stake.

Ultimately several conclusions can be drawn regarding the nature of the regime topic and how conducive it is to regime formation. First, where revenue flows are comparable, the rules allocating taxing rights have minimal distributional effect and are best characterized as a conflict over means. Not surprisingly, such conflicts are, on balance, easier to resolve. Second, where revenue flows are not comparable, actual distributional consequences follow from whether the rules favor the source or residence jurisdiction. One country will end up with more goods (revenue) depending on how the rules are drafted. Not surprisingly, this tension generates more disagreement than the conflict over means. Third, whether the distributional dispute is a conflict over relatively assessed goods and absolutely assessed goods can not be definitely answered, but the unique nature of money suggests that it may be somewhere in the middle. Money is not as benign as wheat on the negotiating table, nor is it as contentious as weapons, and certainly it can be converted into either. Finally, although the analysis has depicted a negotiation in which the flows between country A and country B are either equal or not equal, this picture oversimplifies the negotiation process. It is quite possible that on some issues A and B have comparable flows, but that on others they have different and unequal positions. Thus, within the negotiation over treaty rules there may be some conflicts over means and some conflicts over goods. If the latter predominate, negotiations may be more difficult than if the points of difference between A and B are few and small in value.

iii. Background Factors Influencing Regime Formation

Having now determined the game model, and the probable classification of the conflict, the last basic step in the neoliberal approach to regime analysis examines the “background factors” – the constellation of factors that can influence the success or failure of regime probability of regime formation).
formation. Some of the major factors include frequency of interaction, type of foreign policy practiced by the state, distribution of issue-specific resources, the presence or absence of salient solutions, the number of actors in the issue area, and the availability of leadership. Not all factors are expected to influence all game structures equally. For double taxation, the frequency of interaction combined with the “semi-exogenous” shock of World War I’s high tax rates may have galvanized a measure of interest in resolving double taxation issues. The distribution of issue specific resources is another way of describing whether the two states have comparable investment flows.

As to solutions, the mechanisms for reducing double taxation (credit or exemption) were theoretically well established and relatively clear in their application. The difficulty arose in agreeing to a given application’s distributional effects and coordinating the intersections of the states’ distinctive domestic regimes. The difficulty in this area seems borne out by the fact that although the model treaties were the product of a group of nations (ICC, League of Nations, OEEC, OECD, and U.N.) the end products have been bilateral treaties. This held true, despite the fact that many participants along the way advocated for a multilateral treaty. Because of the large number of actors in the issue, any formal, binding multilateral agreement would have more difficult. Some analysts have argued that the model treaties should be viewed as effectively a multilateral agreement on double taxation with the individually negotiated bilateral treaties serving as the equivalent of national reservations to the basic document. Obviously, this analogy is a stretch but it does capture the sense that there is more multilateral agreement on double taxation than the use of bilateral treaties would otherwise indicate. Moreover, the fact that countries approving the OECD model treaty can make their own individual observations and reservations to the model commentary, supports the image of multilateral agreement with

---

221 See supra text accompanying note ___.

222 Except for the U.S. model, but even this showed the strong influence of the multilateral models that preceded it, and indeed, on which the U.S. model is structurally based.

223 Cite Adams; U.N. Introduction to the Model Convention.
specified national reservations. 224

The final factor, leadership, is both elusive and potentially quite powerful. 225 Certainly in the early years of the U.S. income tax, a few key individuals played a pivotal role in the formation of tax policy including T.S. Adams (who pressed for the enactment of the U.S. foreign tax credit and urged the importance of source jurisdiction) and Seligman (who was one of the “four economists”). 226 The presence of a U.S. leader (Adams) advocating a position on double taxation that “favored” source countries must have diffused some tension between the debtor and creditor nation camps in the 1920s. Of course, the fact that Adams had difficulty persuading the United Kingdom shows limits to leadership power. Furthermore, the differing positions of two major U.S. figures in the 1920s, Adams and Seligman, reveal that even a single country may not have a uniform voice, although Adams clearly prevailed both in terms of domestic legislation (the foreign tax credit) and international support for source jurisdiction. In later years, Stanley Surrey played a dominant role in both the OECD and the U.N., demonstrating the potential for individuals to transcend countries and organizations. 227

The secondary factors, while not dispositive in terms of regime formation nonetheless generally resonated with the actual experience in the double taxation area. Once we have amassed a body of research on international tax case studies from an IR perspective, it will be useful to consider whether particular factors regularly play a larger role in tax regime formation. In other subject areas, such as environmental law, extensive analysis of many case studies has demonstrated the recurrent importance of certain factors in establishing regimes in those fields. 228

224 See, e.g., Commentary to OECD Model Treaty.

225 Cite discussion of environmental case where individual leadership was a predictable important factor in regime formation.

226 See Graetz & O’Hear, supra note ___ at ___. See also Edwin R.A. Seligman, DOUBLE TAXATION AND INTERNATIONAL FISCAL COOPERATION (1928) at 114-165.

227 See supra note ___.

228 See, e.g., Oran R. Young & Gail Osherenko, “Testing Theories of Regime Formation: Findings from a Large Collaborative Research Project,” in REGIME THEORY AND INTERNATIONAL RELATIONS (ed. Volker Rittberger, 1993) (reviewing the findings of a large multi-case study research project in environmental law) at 223;
c. **Cognitive Regime Theory**

Although cognitivism has not stood alone as an independent theory of regime formation, it has developed (in its “weak” theory version) a complementary role with the neoliberalist and neorealist traditions, through its examination of the impact of epistemic communities and the use of knowledge. 229 Even in the absence of a dominant state, regimes can be created when a community of experts active across multiple states develops a vision, develops public awareness, and lobbies the government. Once a determination has been made that a particular regime is best characterized as either neorealist or neoliberalist, the regime is then examined to consider whether epistemic communities significantly impacted regime formation.

In this double taxation case study, several international organizations (League of Nations, OEEC, OECD, and U.N.) played a pivotal role in gathering and organizing expertise, and using it to provide model regimes (the treaty models). In fact, the League of Nations labeled the body it assembled to draft a model as the “Committee of Technical Experts.” 230 The organizations and their “experts” provided a setting in which many of the detailed issues of international taxation could be explored and elaborated by and among those with detailed knowledge. In addition, the momentum within these organizations to identify, enumerate, and solve the problems of double taxation propelled countries toward model treaties. The resulting model treaties themselves adopted an independent life of their own. This latter point can be observed through the U.N.’s decision to closely follow the OECD model and depart only where truly necessary to achieve the specific goals of the U.N. 231 The OECD model had become such an established benchmark that unnecessary departures from its structure and content would have made the U.N. model treaty less attractive and less susceptible to adoption.

---

229 See supra text accompanying notes ___, identifying four basic regime patterns: (1) neorealist, (2) neoliberalist, (3) neorealist impacted by epistemic communities, and (4) neoliberalist impacted by epistemic communities. Epistemic communities are groups of experts in a given field that can shape policy, disseminate information and facilitate agreement. See supra text accompanying notes ____.

230 See Seligman, supra note __ at 143-165.

231 See supra note __.
Although international organizations played a major role in treaty development, they have typically comprised representatives from the individual countries. Differences in opinion can and do exist between the representatives that states send to an organization and the states themselves, but nonetheless, these state based international bodies do not represent the most classic version of an “epistemic community” as understood by neoliberal theory. This tension between the organizations as independent forces of experts and as reflections of state interest can be seen in the observations Seligman made regarding the League of Nations Committee of Technical Experts:

When the economic experts first met, there was form the outset nothing but cordial cooperation, as is entirely natural in the case of those who pursue the career of science and who are interested only in the attainment of truth. When, however, the technical experts came together, their concern was primarily to enter into some arrangement which would be politically agreeable to their respective countries. Everyone accordingly brought with him, together with a desire to arrive at a final arrangement, a feeling we, shall not say of hostility or even suspicion, but at all events of doubt . . . [but] when they learned to know each other more intimately; and especially in proportion as they subjected to the indefinable but friendly atmosphere of the League of Nations, their whole attitude changed. Suspicion was converted into confident; doubt was resolved by the feeling of certainty of accomplishment; and aloofness gave way to warm personal friendship which contributed materially to the smoothing out of difficulties.

The picture conveyed by this quote suggests a rich and complex dynamic within the League of Nations that is not adequately captured by labeling the organization as a collection of countries. Moreover, these organizations have a leadership structure within them, beyond the basic membership through representative countries. For example, the OECD is led by the Secretariat

---

232 This point actually raises a complicated question about who speaks for the “state” and whether it makes sense to describe the state as having a single view. In the most simplistic usage, where the state’s view refers to the bottom line position on an issue, there still can be a gap between that position and the views of the individual representatives to the organization. This is most likely if the individual has represented the state in the organization for a number of years as domestic political power changes.

233 A more classic example would be a group of scientists concerned about a developing environmental hazard on which countries are not currently focused. If the scientists are able to identify and understand the underlying environmental issue and reach a basic consensus on how it should be handled, their views could both force countries to take the issue seriously and shape the direction of the policy they develop.

234 Seligman, supra note __ at 143-144.
General, who leadership is crucial to the organization and must come from a sense that he or she is not “merely” the advocate for particular country.\textsuperscript{235} Even if these organizations are not the prototypical epistemic community, they play a sufficiently similar role to require comparable attention.

4. Conclusions from the Double Taxation Case Study

The application of regime theory analysis to the double taxation case study generates the following specific conclusions (more general observations are discussed in Part III). First, the description of the double taxation relief system as a regime, with the critical components of principle (double taxation is harmful), norm (residence should yield to source) and rules (details coordinating the intersection of two countries tax laws), seems affirmatively accurate and descriptive.

Second, of the two dominant models of regime formation, the neoliberalist most accurately reflects the experience of the double taxation regime. Although the neorealist focus on power (including economic power) may be useful in explaining why one distributive rule prevails over another in treaty negotiations, the neoliberalist model (which looks beyond power to the impact of game theory, issue type, and background factors) offers a more comprehensive understanding of the regime formation process here. For example, it helps explain why countries negotiate these treaties despite the availability of a unilateral solution to double taxation, and why some countries pursue treaties and others do not.

Third, within the neoliberal model, the double taxation regime process most closely mirrors a coordination game. As such, we would expect that agreement would be relatively easier to achieve because there is no monitoring problem (i.e. no need to prevent defection). The primary challenge in a coordination game is the need to reach a decision that may have some distributive effects. The greater the distributional component of the coordination game, the more difficult it is to reach a consensus. Thus, where negotiating countries A and B are both developed countries with similar investment flows, fewer distributional issues should arise. If A is a developed country and B a developing country, the selection of regime rules will carry distributional consequences that will impede agreement.

Fourth, the nature of the conflict in the double taxation regime impacts regime formation. Following the observations articulated above for the game theory aspect of the analysis, where A and B are similarly situated, the conflict can be characterized as one over the “means” of implementing double taxation relief. Regardless of whether the selected rules favor residence or source, A and B will not experience any serious distributional impact. If, however, A is a developed country and B is a developing country, then rules that favor the residence country will typically favor A – adding a distributional dimension to the conflict. Not surprisingly, this factor increases the tension and difficulty in reaching agreement on the rules.

Fifth, the background elements (including frequency of interaction, availability of salient solutions, impact of exogenous shocks, and presence of strong individual leadership) enhance the explanatory success of the other factors. For example, influential and credible individual leaders (Adams, Seligman, and Surrey) appear prominently in the double taxation story. Is this characteristic of regime formation in the tax area? If it is, or is important in a subset of cases, how might that realization impact regime strategies?

Finally, neoliberalist regime theory alone can not adequately account for the double taxation regime. The epistemic community (as described by cognitivist theory) served as a driving force, both in terms of providing a forum for discussion and providing a base of expertise to structure the debate. Although the precise contours of this epistemic community and its role merit further attention, its sustained importance in the process is clearly demonstrated over the decades.

III. DEVELOPING THE IR-INTERNATIONAL TAX RESEARCH AGENDA

A. Introduction

The double taxation case study demonstrates the role that IR theory and methodology can play in furthering our understanding of even the most familiar of international tax stories. A recurrent question in international tax is how and whether countries can reach agreement on some problem or set of issues. Whether the topic is transfer pricing, documentation standards, withholding or arbitrage – the real question has been whether a meaningful regime can be
formed. The IR regime theory literature is devoted to understanding and answering that question. Based on the initial application of regime theory to the double taxation case study, (1) What broader observations can we make, and (2) How should we design a research agenda to further the development of IR regime theory in international tax?

B. General Observations

1. Unique Role of Taxation: The first observation grows from the fact that regime theory has been applied to many fields other than tax (e.g., environmental, security). How is international tax different from these other areas and why might that matter to regime theory? In many other fields such as environmental or communication law, the government intervenes (i.e. regulates) because the market does not function adequately on its own. For example, environmental regulations often address the reality that certain behaviors have environmental costs that are not borne by the parties engaging in the conduct. Such externalities are beyond the scope of the market to address and require governmental intervention. In contrast, governments issue tax regulations for distributive/revenue reasons, rather than to remedy market failure. In this way, states have a direct and unique interest in their role as tax regulators.

In shifting to the international sphere, the government’s role in providing tax rules for cross border transactions is no different. The state, as part of its “domestic” revenue collection function, must define what is income and who must pay tax. Traditionally, states have taxed foreign source income earned by their residents as well as income earned in the state by nonresidents. Many countries are engaged in this taxing function simultaneously and not surprisingly clashes occur and cooperation is required to resolve the conflict—they need a regime. As countries form a regime, such as the double taxation regime, they must integrate their multiple roles. The states can be seen as actors pursuing an activity (taxation) and the international regime (in place of a supranational government) regulates the activity so that it can

\[236\] Tax, as a subject, is a complete construct of the government. Tax issues would not exist in the absence of government tax regulation (although distributional needs would). And the type of tax regulation is not necessarily related to the particular distributive/revenue goal being met. Although the primary force behind tax law is the need to collect revenue, governments do use the tax system for secondary purposes as well, including the promotion of social, political or economic goals.
take place more efficiently (e.g., reduced administrative burdens on parties, information sharing, or increased efficiency of cross border investment). This characterization of regimes in international tax supports the view that neoliberal regime theory better describes regime formation in tax. Certainly that is the case for double taxation; further research will establish whether the majority of tax regimes share that quality.

In addition to improving taxing efficiency (for example helping a state collect tax from taxpayers duly owing tax or ensuring that tax rules do not unnecessarily impede business), the rules in a tax regime can allocate revenue among the competing states. Several of the double taxation regime scenarios (e.g., where A was a developed country and B was a developing country) included some allocation of tax revenue that had distributional effects (reflecting the limited pool of tax dollars). Governments are real stakeholders as to this revenue and have expressed a strong sensitivity to incursions against their “sovereignty” in the tax field. Whether this sovereignty claim can be justified, it may relate to the state’s somewhat special role as a different kind of stakeholder in tax matters.

2. Importance of the Identity of Regime Participants: Of course countries are not equal in terms of their power or their resources relevant to the regime area. In the double tax regime this point emerged through the recognition that some double tax rules would have distributional effects where the flows between countries were not equal (as with a developed/developing country pair). However, even all developing (or all developed) countries are not equal. The game theory analysis demonstrated that: (1) a developed country pair could have distributional concerns in negotiating the double taxation rules; (2) a single negotiation between two countries could have include some rules with no distributional effects, and other rules with a small or large distributional effect; and (3) some developing countries will pursue a double taxation regime while others will not, depending in part on the type, nature, and quantity of foreign investment into the developing country and the strength of the developing country reputation as an investment destination.

3. Details Matter: A productive and accurate use of the game theory models requires a detailed knowledge of the substantive area being evaluated. In the double taxation case study scenario
using a developed/developing country pair, what initially looked like a suasion game (in which A, the developed country must “pay” B to participate) looked on further reflection like a coordination game because both A and B want to reach agreement on double taxation to facilitate investment and neither benefits from defection.

4. Benefits of Game Theory Analysis: Evaluating regime formation from a game theory perspective offers a narrower and a broader benefit. Narrower: Using game theory requires the careful examination of parties, goals, facts, constraints and structure in a precise way that disciplines one’s analysis of the case study. A negotiator or other participant in an international regime that undertook a game theory analysis would emerge a much better informed player in the process. Broader: To the extent that evidence suggests certain game types are more susceptible to agreement or regime formation, this knowledge may offer some predictive guidance on forward looking inquiries.

5. Regime formation and path dependence: When states gather to evaluate a problem and seek to establish a regime, they are likely to approach the problem from their own perspective. As a result, participation in the initial formation of the regime can be critical to shaping the debate and its lasting consequences. The “staying power” of regimes, exacerbates this “first mover” problem. If an organization (or set of states) makes the first move at resolving an issue (even if that agreement fails to reflect the interests of all possible participants), then the regime, once established, may have a life of its own that effectively constrains the ability of a second generation agreement from gaining the same degree of prominence.

C. Research Agenda

How should we design a research agenda to build upon the application of IR theory to international tax? What questions are especially pressing? Going forward several avenues of inquiry are particularly important. First, as has been done in other substantive fields, it will be valuable to develop a stable of case studies to help identify common issues, patterns and

---

237 See, e.g., Hasenclever, supra note __ at 38.

238 Note for example the use of the same treaty structure and framework by the various double taxation model treaties. This was true even in the case of the U.N. (in drafting its model double taxation treaty)
problems and serve as a data base for testing various aspects of regime theory. It may be possible to discern particular factors central to regime formation in the tax area, mirroring comprehensive efforts in the application of regime theory to environmental law.\footnote{239} Included in the group of new case studies should be example of “complete” failure to form a regime as well as examples in which the states were partially successful in establishing a regime (perhaps by changing or limiting the parameters of their negotiations). Candidates for study include the PATA\footnote{240} agreement on documentation for transfer pricing,\footnote{241} Advance Pricing Agreement Programs,\footnote{242} the European Union savings directive,\footnote{243} and efforts to limit “harmful” tax competition.\footnote{244}

Second, the role of epistemic communities in taxation requires careful examination. Rather than approach this question solely through the case studies suggested above, it would be useful to directly study the major groups involved in international tax policy, including the League of Nations, the OEEC, the OECD, the U.N., and the International Fiscal Association. What role have they played in identifying and clarifying issues, directing discussion, and shaping which specifically sought to counter the OECD model’s focus on the needs of developed countries.

\footnote{239}{Young & Osherenko, \textit{supra} note __.}

\footnote{240}{PATA is the Pacific Association of Tax Administrators.}

\footnote{241}{PATA issued agreed documentation guidelines in March 2003. The introduction to the guidelines states that the PATA members “are providing principles under which taxpayers can create uniform transfer pricing documentation . . . so that one set of documentation can meet their respective transfer pricing documentation provisions.” PATA, Introduction to Transfer Pricing Documentation Package (2003).}

\footnote{242}{The Advance Pricing Agreement Program (“APA” program) allows taxpayers and governments to agree on the pricing and allocation of profits from related party transactions in advance of the transactions. Following the introduction of the APA program in the United States in the early 1990s, many other countries have formally or informally instituted such programs as well. \textit{See} Diane M. Ring, \textit{“On the Frontier of Procedural Innovation: Advance Pricing Agreements and the Struggle to Allocate Income For Cross Border Taxation,”} 21 \textit{Mich. J. of Int’l L.} 143, 163-169 (2001).}

\footnote{243}{The savings directive concerns the treatment of cross border payments of interest within the European Union. \textit{See}, \textit{e.g.}, George Guttman, \textit{“EU Taxation of Foreign Interest Is A Multiple-Choice Game: Withhold, Report, Or Ignore,”} 28 \textit{Tax Notes Int’l} 459 (Nov. 4, 2002).}

\footnote{244}{Beginning with the issuance of its 1998 Report on Harmful Tax Competition, the OECD has sought to mobilize support for constraints on certain types of tax competition. mobilize support for constraints on}
conclusions? In looking for their influence, we should go beyond their direct roles in producing agreements, to consider more subtle contributions. For example, long before an agreement is on the table, do these epistemic communities raise the profile of issues by generating discussion and study? Do they increase the general comfort level on various rules and outcomes by presenting and reviewing them repeatedly and extensively? Based on the double taxation case study, organizations that are part of the epistemic community are worth our attention, are worth joining, and are a good vehicle for advancing issues and ideas about which you have strong views.

On a related point the impact of individuals on regime formation in taxation must be examined. Influential individuals may be acting inside and or outside of an epistemic community or organization. This issue can be approached in two directions: (1) through direct study of critical figures in international tax (i.e. mini biographies), and (2) through an awareness of any individual actors in the case studies undertaken (e.g., tax competition).

Finally two connected issues that have only been hinted at so far deserve analysis – the impact of the domestic tax and political environment on formation of regimes, and the role of sovereignty in regime formation and regime failure. Asking about the domestic environment acknowledges the reality that states are not monoliths and may not speak with a single voice or follow a single vision. (In a limited way this occurred in the tension between the views espoused by Adams and by Seligman on double taxation). States are not merely actors on the international stage; they are entities responsive to and in many cases shaped by the domestic sphere in which they operate. One of the many threads that surfaces in the domestic political rhetoric directed at international tax matters is the question of sovereignty and the degree to which any proposed regime would and could impinge on the nation’s sovereignty in tax matters. The examination of “sovereignty” in regime formation may be connected to this domestic side of regime formation as well. In some domestic debates on international taxation (and on international agreements generally) sovereignty serves as the focal point for the discussion.

CONCLUSION

As nations devote more attention to the interaction and coordination of their activities in a global community, including international taxation, our understanding of how these efforts are structured, shaped, and influenced becomes critical. Regime theory from the IR literature offers an important framework for examining the agreement process between and among nations. Through analysis of international tax case studies, we can identify common themes in the regime experience in international tax that may be generalized and may enable us to predict where and when regime formation efforts are likely to be successful and how that success can be fostered. This work is not only the domain of tax scholars. Just as researchers in IR have developed extensive case studies in other substantive fields, they should now turn to taxation. This expansion not only offers a new realm in which to test and apply their theories, it also offers an opportunity to examine regime formation in a context in which the states have a special role (revenue collector). Ultimately, work from researchers in both taxation and IR theory should together guide states in their efforts to design and implement international tax policies.
Taxing Royalty Payments with a Developing Country Source: A Comparison of Canada and Australia’s Tax Treaties (3/23/2006)

By Kim Brooks, UBC Faculty of Law
I. The Growing Significance of Technology Transfer and the Consequent Importance of Taxation at Source

It is a widely recognized norm of international tax that countries have the jurisdiction to tax individuals or entities that are resident in their jurisdiction on their world-wide income. So, for example, Canada maintains the right to tax an individual or company resident in Canada on all of his or its’ income, whether earned in Canada, Australia, or Turkey. At the same time, it is also recognized that countries have the jurisdiction to impose tax on all income that has a source within its jurisdiction, even if the owner is not resident there. So, for example, Canada asserts the right to tax residents of other countries on the income they earn with some nexus, or source, in Canada. To illustrate, a French resident, who has no personal physical presence in Canada, but who owns land in Canada that she sells, may be subject to tax in Canada on the gain associated with that sale. Obviously, this gives rise to the well-known problem of international double taxation. If a person or corporation resident in one country earns income in another country, she or it will be taxed on that income in both the country of residence and the country of source.

The well-accepted, unilateral way of avoiding international double tax is for the residence country to provide a tax credit for the taxes paid in the source country. To rely on the France example again, to unilaterally avoid double taxation, the French government would give the taxpayer credit for the Canadian taxes paid, hence reducing the French taxes owing by the amount of Canadian tax has been paid and relieving the effect of the imposition of tax by two jurisdictions. That is, in the unilateral relief example, the source country (in this example, Canada) has the first claim to tax income with a source in its jurisdiction.

Assuming that countries could agree on where taxpayers were resident and in what jurisdiction income was earned, this simple system could avoid all double taxation problems. Indeed, at the turn of the 20th century as the various industrialized countries introduced income taxes, this is how the double tax problem was resolved for transnational income flows. However, particularly in the 1920s, because of the disparate rules between countries relating to residence and source, and because the possibility of multiple taxation was seen as hindering international trade and commerce, as international commerce increased, the industrialized countries began considering the adoption of bilateral tax treaties to prevent double taxation.¹ The simplest way to achieve the avoidance of double tax was to enter into an agreement that essentially limited the source countries’ ability to tax. Provided both countries had about equal flows of

¹ See, for example, the description of the purposes of tax treaties in Stanley S. Surrey, “United Nations Group of Experts and the Guidelines for Tax Treaties Between Developed and Developing Countries” (1978) 19 Harv. Int’l L. J. 1 at 3. (“Tax treaties are bilateral conventions negotiated between countries for the primary purpose of resolving double taxation problems which arise from the assertion by more than one country, through their internal laws, of jurisdiction to tax the same item of income.”)
income, this method – privileging residence taxation - would not result in any loss of revenue to either country. Many of the very early tax treaties were entered into between developed nations and allocating the ability to tax to residence countries both worked well, and was relatively easy. In fact, in 1963, the Organization for Economic Cooperation and Development, or OECD, developed a model treaty that reflected this arrangement and, for most kinds of income, it allocated taxing rights to the country of residence.2

Where, however, flows of income are not equal, restricting source country taxation reduces the revenues of source countries and increases the revenues of residence countries.3 Therefore, as developing countries began to enter into tax treaties, there was increased attention placed on the allocative consequences of privileging residence-based taxation. While the model treaty drafted by the OECD might have been fair when entered into by two developed countries, it was seen by many developing countries to be decidedly unfair when it was entered into between a developed (capital exporting) country and a developing (capital importing) country. In spite of this, most developed countries insisted that if developing countries wished to enter into tax treaties with them those treaties would be based on the OECD model.

Developing countries, as net capital importers, recognizing that their aims in entering into tax treaties were perhaps different than the aims of two countries with reciprocal trade flows, attempted to resist the use of the OECD model treaty. In 1967, the Economic and Social Council of the United Nations established the United Nations Group of Experts on Tax Treaties Between Developed and Developing Countries. This group of experts met a number of times between 1968 and the release of an alternative model tax convention. The UN model convention was adopted by the U.N. in 1980.4 The UN treaty, although it is based to a surprising extent on the OECD model, is more favourable to developing nations than its OECD counterpart. Generally stated, then, the goal of the UN model treaty was to leave significantly more scope for developing countries to tax income based on its source than permitted under the OECD model.5

Therefore, while tax treaties are most frequently lauded as a mechanism for reducing double tax and for facilitating international trade and investment,6 for developing countries their most important function may well be the degree to which they allocate tax revenues between the countries that have entered into the bilateral agreement. Tax treaties do not generate tax revenues as a whole, of course; instead, they simply allocate

2 Cite to OECD model.
3 Stanley Surrey describes this as a compromise in the allocation of taxing jurisdiction: “The OECD Model rests essentially on two propositions: (1) the country of residence will eliminate double taxation through a foreign tax credit mechanism or through exemption for foreign income from tax; and (2) in turn, the country of source will considerably reduce both the scope of its jurisdiction to tax at source and the rates of tax where jurisdiction is retained.” (at 8).
4 Cite to UN convention.
5 See Surrey, supra note ? at 10. (“The basic departure from the OECD Model is, thus, the emphasis placed by the U.N. Group on taxation at source.”)
6 This is apparent even in reviewing the title section of Canada’s tax treaties, which often says [complete with how often avoid double tax/promote investment.]
those revenues between the parties to them. So, for example, assume that a Canadian resident earns $100 in royalty profits. If Canada taxes royalty income earned worldwide by its’ residents at a rate of 40%, and if the country of source exempts the royalty income from tax altogether, the Canadian government would exact $40 in tax from its’ resident. If, however, the source country imposed a rate of tax on the royalty payment of 40%, the Canadian resident would pay $40 to the source country. Normally, assuming that the foreign tax credit mechanism worked perfectly and that the rate of tax at source did not exceed the residence rate of tax, the Canadian resident would then pay no taxes to the Canadian government on the royalty income. The Canadian resident may have a preference for one system over the other if she has a preference for her tax dollars to be collected by the source country as opposed to the residence country or vice versa, but presumably this system leaves the Canadian resident indifferent from an economic perspective on where she earns the royalty income. No matter where it is earned, $40 of tax is paid.

Although, based on this illustration, the Canadian resident may be indifferent about the tax allocation system since the total taxes she pays are the same regardless of who collects them, the two governments to the treaty may care a great deal about who gets to collect the taxes. Each government may have important uses for the revenue to be collected, may have different administrative mechanisms and abilities to actually collect those revenues, and may wish to deliver particular incentives through the tax system that might ultimately influence the rate of tax they would like to impose on the income. In other words, the tax system might be intimately connected with each country’s plans for economic development, poverty alleviation, social infrastructure development, and aid programs.

This chapter argues that developed countries like Canada ought to use their tax treaties to facilitate the allocation of a fairer share of income to developing countries. This argument is made in the context of the taxation of royalty payments, which present one of the most extreme examples of where privileging residence-based taxation generates unfairness for source countries. This is the case because, unlike other forms of property income like dividends or interest, the OECD model convention provides that royalty payments are not subject to withholding tax at source. They are only taxed in the country of residence. Although the privileging of residence-based taxation of royalty payments may always have been unfair to source countries, that unfairness likely has been exacerbated by two trends over the last forty years.

First, the value of royalty payments has presumably grown significantly over the past forty years. Yet, despite increasingly sophisticated approaches valuing royalty payments, the upsurge in reliance on outsourcing (and the related transfer of technical knowledge), the dramatic increase in the use of computers, computer processes and software, which need to be transferred to almost any jurisdiction where a company carries on operations, and the increased ease with which intangible property giving rise to royalty payments may be relocated anywhere in the world the OECD has not amended its position on the imposition of a royalty withholding tax. Although in 1963 the failure to grant source jurisdictions with the right to tax royalty payments may not have been so egregious, and
may have been a sensible concession to administrative considerations, given these
dramatic changes, the value of royalty payments and the number of property transfers
giving rise to royalty payments has presumably increased significantly, leading to a much
increased loss of tax revenues for source-countries. Put simply, property that gives rise to
royalty payments is valued using increasingly sophisticated methods (and is worth more)
now than it was forty years ago, and the number of cross-border transactions giving rise
to royalty payments has increased. The revenue implications of the early administrative
concessions should be seriously questioned, and the attachment to residence-only
taxation of royalty payments should be abandoned.

Second, an increasing number of transactions might be characterized as royalty
payments. Despite the explosion of e-commerce, royalties have been defined in the
OECD model convention in roughly the same way for over forty years. A royalty is
defined to mean

...payments of any kind received as a consideration for the use of, or the right to use, any
copyright of literary, artistic or scientific work, including cinematograph films, any patent, trade
mark, design or model, plan, secret formula or process, or for information concerning industrial,
commercial or scientific experience.\footnote{Insert cite to OECD model}

Although this definition has remained virtually unchanged, the trade between nations in
the kinds of property that gives rise to royalty payments has changed immensely in the
same period. While in the past, sales of books, for example, into a jurisdiction may have
easily been characterized as giving rise to income from business, it is more difficult to
classify income from the transfer of the same book electronically over the Internet.
These characterization issues have plagued the OECD and commentators, who have
struggled with drawing clear lines between business income and royalty payments. The
difficulty of the line-drawing exercise is not disputed here, it is simply noted as evidence
that transactions which historically would have given rise to income easily defined as
income, now may give rise to income better characterized as royalty income.

This chapter proceeds as follows. Part II reviews the role and effect of the royalties
article in tax treaties. It focuses on three aspects of the royalties article. First, it discusses
the importance of taxing royalties at source. I argue in particular that developing
countries should be granted the right to tax royalties at source at rates that would give rise
to tax revenue that might be used to further their domestic development goals. Second, it
both reviews the meaning given to the definition of “royalty” in the royalties article and
describes the privileging of the application of the royalty provision as opposed to the
business profits article. I argue that despite the difficulties inherent in distinguishing
income from business from income from property (here, royalties), that distinction should
be maintained. Third, it examines the exclusion of particular types of payments from the
royalties article. Here, I argue that exemptions from royalty taxation should be
eradicated. In Part III, Canada and Australia’s tax treaties are canvassed to assess the
positions those jurisdictions have taken on the taxation of royalty payments. This
assessment reveals that Canada is both more likely to have negotiated a higher
withholding tax rate or non-reciprocal withholding tax rates, and that Canada is more likely to use the royalties article to privilege residence taxation of particular kinds of royalties. Finally, Part IV concludes by summarizing the modest recommendations suggested in this paper.

II. Allocating Revenues Through the Royalty Article in Tax Treaties

The royalties article of a tax treaty is commonly drafted to serve a number of purposes. It may set out the jurisdiction of the treaty partners to tax based on either source or residence, or both;\(^8\) where source jurisdiction is granted, it will impose an applicable withholding tax rate;\(^9\) it may provide guidance about the geographical location where a royalty arises;\(^10\) it may provide some scope to the meaning of the word royalties by setting out payments that are expressly included in the article;\(^11\) it may privilege the application of the business profits article where the royalties arise through a permanent establishment and the right or property that gives rise to the royalty is effectively connected to the permanent establishment;\(^12\) and, finally, it may permit tax authorities to alter the amount of the royalty payment where the payment is made between parties with a special relationship and it exceeds the amount that would have been agreed to by parties without such a relationship.\(^13\) Although the royalties article serves a number of functions, the three functions of particular importance for developing countries are the grant of jurisdiction to tax at source, any articulation of how royalties might be distinguished from other types of income, and any exclusions from the royalties provision. This part of the article emphasizes the importance of permitting taxation of royalties at source, details the significance of maintaining the distinction between business profits and royalties, and advocates for the removal of exemptions from the imposition of withholding tax at source.

A. Taxing Royalty Payments at Source

Source states have strong economic claims that justify the taxation of royalty income at source. Despite the strength of these claims, the most significant difference between the UN and OECD model conventions is their position on the ability of source jurisdictions to tax royalties. As noted above, although the OECD model convention provides for both residence and source taxation when property income in the form of interest or dividends is paid, for royalty payments only the residence country is granted jurisdiction to tax the royalty.\(^14\) The UN model, in contrast, permits both residence and source-based taxation.\(^15\) Developed countries entering into tax treaties with developing countries

\(^8\) For example, see OECD model convention Article 12.1; UN Model convention Article 12.1 and 12.2.
\(^9\) UN model convention Article 12.2.
\(^10\) UN model convention Article 12.5.
\(^11\) OECD model convention Article 12.2; UN model convention Article 12.3.
\(^12\) OECD model convention Article 12.3; UN model convention Article 12.4 (which additionally includes a force of attraction requirement).
\(^13\) OECD model convention Article 12.4; UN model convention Article 12.6.
\(^14\) cite article and section.
\(^15\) Cite article and section.
should follow the UN model convention and should therefore permit the taxation of
royalty payments at source.

The arguments marshalled in favour of an exclusive residence-based taxation of royalty
payments are insufficient to justify depriving developing countries of a share of the tax
revenues associated with the use of property giving rise to royalty income in the
developing country. This part of the paper reviews each of the claims in favour of
exclusive residence-based taxation of royalty income and shows each to be wanting.

First, it might be argued that the residence country has a more principled claim to tax the
income from royalties since the owners of the intangibles have invested the time and
expense of developing the intangibles and that, therefore, the residence jurisdiction has a
much closer economic nexus to the activity that gave rise to the value of the property than
the source country does. However, at best this is a sufficient justification for preserving
some ability of the residence country to tax royalties, and not a justification for exclusive
jurisdiction. The source country has a similar claim to a significant economic connection
to the revenues produced through the use of the property giving rise to the royalty since
the income arose in the source jurisdiction.\(^{16}\) In addition, where the property that gives
rise to the income is an intangible, presumably the intellectual property protections
provided by the source country dwarf in significance those provided by the residence
state. In other words, if the source state did not offer a legislative and enforcement
regime that adequately protected the intellectual property rights of the taxpayer engaged
in the activity in the source state, there would be little income to earn from that property
in the long run. Therefore, requiring the source country to abandon its ability to tax
income derived from a royalty payment unduly benefits the economic connection of the
residence jurisdiction without meaningful justification for doing so.

Second, it might be argued that the country of residence is better able to design a tax
regime that accurately reflects the expenses associated with the production of intangibles
and other property that gives rise to royalties and therefore is better able to tax the real
income associated with the use of those intangibles. Since royalty payments are taxed at
source through the imposition of a withholding tax, no credit is afforded for the expenses
associated with the production of the property giving rise to the royalty. However, this is
not an argument in favour of non-taxation at source. Instead, it is an argument that
supports thoughtful negotiation of an appropriate withholding tax rate.\(^{17}\) Indeed, if

\(^{16}\) As compellingly phrased by the four economists who produced a report for the League of Nations on
income characterization in 1923, “The oranges upon the trees in California are not acquired wealth until
they are picked, and not even at that stage until they are packed, and not even at that stage until they are
transported to the place where demand exists and until they are put where the consumer can use them.”
Financial Committee by Professors Bruins, Einaudi, Seligman and Sir Josiah Stamp*, League of Nations

\(^{17}\) In fact, the importance of considering a range of factors, including the expenses associated with the
production of property giving rise to royalty payments, was recognized by the Group of Experts. As a
consequence, the Group of Experts have not set exemplar rates of withholding tax for any of dividend,
interest, or royalty payments. Instead, the applicable rate of withholding is to be left to the negotiators of
the particular tax treaties. Negotiators of tax treaties are urged to consider, among other things, the
expenses that may be associated with the royalty, the costs associated with developing the property, the
developed countries were committed to supporting the revenue-raising goals of
developed countries, and were also committed to ensuring the correct taxation of
income.\textsuperscript{18} developed countries could undertake to refund any overpayment of tax in the
source country to residence taxpayers. This refund would amount to a tax expenditure by
developed countries, designed to support investment in developing countries. In addition,
interest and dividend payments are also subject to withholding tax. It is not clear why the
recognition of expenses supports a position of no source-based taxation of royalty
payments at all, and yet the difficulty of recognizing expenses associated with the
production of interest and dividend payments is not seen to be similarly intractable.\textsuperscript{19}
Lastly, a withholding tax, which does not account for the expenses associated with the
production of the royalty return, may be justified on the basis that, for the most part,
intangible property like patents and processes are licensed to developing countries only
after their value has been fully recouped in developed countries, and therefore there is
little real expense associated with the use of the property in the developing country.\textsuperscript{20}

Third, exclusive residence-based taxation might be justified on the ground of
administrative ease. These administrative justifications for residence-only taxation might
be usefully divided as falling into three categories: first, arguments based on the
difficulty of determining the geographical location of the property that gives rise to the
royalty; second, arguments based on the high level of tax avoidance that might arise as a
consequence of the ease with which intangibles are (re)located; and, third, arguments
based on the inability of source countries to administer a withholding tax on royalty
payments.

Determining the geographical source of royalties associated with intangibles, in
particular, is notoriously difficult – their location is easily moved, and they are frequently
costlessly shared, for example. These difficulties have been exacerbated by the
development of ecommerce, and electronic communications, which mean that over time
larger numbers of intangible assets have been produced. While determining the
geographical source of a royalty payment is difficult, the problem is not intractable.\textsuperscript{21}

\textsuperscript{18} By “correct taxation” of income, I mean the taxation of income that permits the deduction of all
appropriate expenses in calculating profit.

\textsuperscript{19} This is not to deny the unique difficulties presented by royalties related to intangible assets in particular.
The creation of intangibles may involve extensive expenses on failed experiments, for example. Should
these costs be considered in attempting to estimate roughly the costs associated with any particular royalty
revenue stream? If so, how? Also, in many cases, once a particular intangible property has been created, it
can be used costlessly by many – so, if a particular technology is employed in the developed country, and
can be used costlessly in the developing country, should any expenses be allocated to the developing
country use?

\textsuperscript{20} This argument is reflected in paragraph 6 of the commentary to the UN model convention. Of course, as
reflected in paragraph 7 of the UN model convention, developed countries reply that it is not true that
valuable intangibles are transferred to developing countries only after they have been fully exploited by
developed ones. See [insert reference to UN model commentary].

\textsuperscript{21} See, for example, suggestions made by [insert.]
In addition, there can be little doubt that the difficulties of determining the geographic location of intangible property, in particular, have created previously non-existent avoidance and evasion opportunities for taxpayers; however, source-based taxation may assist in combating abusive schemes. For example, source-based taxation means that two jurisdictions’ taxing authorities have an incentive to attempt to track and tax the transfer of property giving rise to royalty payments. It also increases the value to developing (or source countries) of entering into robust exchange of information agreements with developed countries.

Finally, the non-taxation of royalty payments at source may be justified on the grounds that developing countries simply do not have the tax administration strength to appropriately enforce and collect the withholding tax. At least one side effect, however, of joint taxation by source and residence countries might be increased communication between taxing authorities, which may assist with the transfer of tax administration knowledge and experience from developed to developing countries and further facilitate strong tax administrations in countries with a high need for government revenues to ameliorate in some cases devastating domestic poverty and other forms of social and economic inequality.\(^\text{22}\) Also, if developing countries were serious about allocating some of the revenues generated from the payment of royalties with a geographical source in a developing country to their residents, they could collect the share of the source tax on behalf of the developing country. This may be done by exemption royalties from withholding tax at source, but by setting a revenue-split percentage rate on the profits collected by developed countries on royalty income earned by their residents in another jurisdiction.\(^\text{23}\)

To summarize, none of the arguments in favour of non-taxation of royalties at source are sufficient to justify depriving developing countries of the revenue associated with the taxation of royalty income. Indeed, there are strong arguments based on the degree of economic connection to the source country, and based on the potential for residence and source countries to work together to combat tax avoidance and evasion to support the continued taxation of royalty payments at source.

**B. Assisting with Characterizing Payments**

One of the most difficult characterization issues, particularly in the context of the transfer of intangible property, is distinguishing between income properly characterized as

\(^{22}\) On the issue of the dangers of knowledge transfer in tax from developed to developing countries see Miranda Stewart, “Global Tax Trajectories”? (complete cite).

\(^{23}\) As a model, this might be designed in the same way that the federal government collects the provincial tax of some provinces. The downside is that the source country would not get to define what constitutes taxable income from property giving rise to a royalty payment and would simply need to accept the domestic rules that applied in the residence country. However, this approach to revenue sharing would alleviate the administrative burden on underdeveloped tax administrations, while still ensuring an equitable share of the tax revenue from royalties was allocated to developing countries. It would make sense under this model for a developed country to enter into these kinds of agreements with only developing countries, since it seems reasonable to assume that developed countries have relatively equal trade flows.
business profits from income properly characterized as a royalty. To illustrate, computer software has posed particularly difficult problems for tax administrators. Part of the problem of characterizing software lies in various ways that it might be packaged and transferred. For example, software may be used to guide the operation processes of the computer itself, or it may be used by end users as a separate application from the computer itself. Computer software can be downloaded over the Internet, or sold on a disk. A program may be made for just one user, or may be sold widely. The rights to the software may be alienated entirely when the software is sold, or the use of the software may be restricted. Both the OECD and UN commentary provide some guidance on how and when payments for the transfer of software should be considered to be subject to the royalties provision of tax treaties rather than the business profits provision.

The royalties and business profits articles of both the OECD and UN model treaties provide some guidance on when one of those articles should apply in lieu of the other. For example, Article 7(7) provides that where a payment might be characterized as a royalty, the royalty article should prevail; however, Article 12(3) provides that where the property that gives rise to the royalties is effectively connected to a permanent establishment, then the business profits article should prevail. The characterization of the payment as appropriately falling within the scope of the business profits article as opposed to the royalties article matters because the tax treatment of those two types of payments differs significantly under the tax treaties. Most importantly, perhaps, business profits are only subject to tax when a non-resident taxpayer has a permanent establishment in the jurisdiction, and then the income earned is subject to tax on a net basis. In contrast, the non-resident taxpayer does not need a permanent establishment in a jurisdiction to be subject to a gross withholding tax on royalties earned in that source country.

The most fundamental issue arising from the characterization of income as business or royalty income for developing countries is whether the royalties article should simply be

---

24 At the 2005 International Fiscal Association Congress in Buenos Aires several panel members discussed whether particular payments in four different cases would appropriately be considered to be “royalties” given the definition of royalties in the OECD model. In each of those cases the panel members differed about whether particular payments were properly royalties, business profits, capital gains, or payments for services. This discussion turned largely on the meaning given to the word “royalties” in the OECD model treaty. See Catherine Bobbett and John Avery Jones, “The Treaty Definition of Royalties” 60(1) Bulletin for International Taxation 23 – 28. In this part of the paper, I am focused not so much on the ambiguities that arise as a consequence of various possible interpretations of the section of the royalties article that gives meaning to what is a royalty as I am on the more general question of what ought to be considered a royalty payment as opposed to business profits and whether royalty payments might properly be excluded from tax in the country of source.

25 Insert cite to the OECD & UN commentary.

26 There are ordering rules for the characterization of a payment as a business profit as opposed to other types of income, for example dividends, interest or immovable property as well. For a detailed review of the ordering rules and their (non)sensibility in the context of the different characterization of income from immovable property and business profits see Brian Arnold’s aptly titled (given that immovable property is set out in Article 6 and business profits in Article 7), “At Sixes and Sevens: The Relationship between the Taxation of Business Profits and Income from Immovable Property under Tax Treaties” (2006) 60(1) Bulletin for International Taxation 5 – 18.
subsumed altogether into the business profits article. Given the differential treatment of royalty payments and business profits, this type of proposal would have significant effects for developing countries. Because business profits are only subject to tax at source once the permanent establishment threshold has been met, royalty payments (characterized as business profits) would no longer be subject to tax unless this threshold requirement, however defined, was met. Most notably, this approach will result in a shift of income from developing to developed countries.

There are a range of arguments against subsuming the taxation of royalties within the scope of the business profits article. Not the least of these responses might be the general critique that the threshold requirement for taxation of business profits, that the business be carried on through a permanent establishment, is outdated and fails to respond sufficiently to the challenges of electronic commerce. This argument presents unique problems for developing countries, who might quite rightly point out that the permanent establishment threshold, even as set in the UN model convention, is still too high to result in the taxation of much business activity, particularly given that a good deal of business activity can now be conducted with only minimal physical presence in a jurisdiction.

Perhaps the strongest argument in favour of requiring a minimum presence in a jurisdiction before subjected royalty payments to tax, whether in the royalty articles

---

27 For a discussion of the issues that arise in characterizing income in the e-commerce context as business income or royalty income, see Jinyin Li, *International Taxation in the Age of Electronic Commerce: A Comparative Study* (Toronto: Canadian Tax Foundation, 2003) at 420 – 444 (Li argues in favour of reducing income to two broad categories – business profits and other income – but retains the distinction between business and royalty income (as a form of other income).); Niv Tadmore, “Further Discussion on Income Characterization” (2004) 52(1) *Canadian Tax Journal* 124 – 140 (Tadmore raises some of the challenges of maintaining the distinction between business profits and royalty income, but does not suggest an alternative to that characterization.). For a discussion of the difficulties of line drawing between business and investment income more generally see, for example, John F. Avery Jones et al., “Treaty Conflicts in Categorizing Income as Business Profits: Differences in Approach Between Common Law and Civil Law Countries” *The Taxation of Business Profits Under Tax Treaties*, Brian Arnold et. al. eds (Canadian Tax Foundation, 2003), chapter 2.

28 Brian Arnold argues in favour of the maintenance of threshold requirements for business profits taxation. He recognizes that these thresholds reduce the income that developing countries retain, but suggests that enforcement difficulties of eliminating threshold requirements for taxation of business profits outweigh the advantages for developing countries. Brian Arnold, “Threshold Requirements for Taxing Business Profits Under Tax Treaties” in *The Taxation of Business Profits Under Tax Treaties*, Brian Arnold et. al. eds. (Canadian Tax Foundation, 2003) at 79. (“One of the effects of a threshold requirement in a tax treaty is to allocate the tax revenues on business profits between the two countries. Arguably, this is not the function of a threshold requirement. As discussed above, the function of a threshold requirement is to provide increased certainty about when source-country taxation applies, to reduce the burden of compliance with source-country taxation, and to ensure the effective enforcement of source-country tax. Tax revenues would be allocated between source and residence countries even if no threshold requirement applied and each country taxed non-residents on all income derived from sources in the country. However, the effect of any threshold requirement for source-country taxation is the allocation of tax revenue on income derived in the source country to the residence country; the higher the threshold is, the more tax revenue will be allocated to the residence country.”)

themselves or under the business profits article, is administrative. It is difficult to enforce tax rules that apply to income earned by a non-resident without a significant physical connection to a jurisdiction. However, as noted above this problem is not intractable.

Before concluding this section on characterization issues, it may be worth noting that while the primarily concerns about characterization tend to be discussed in the context of the distinction between business income and royalties, the non-taxation of royalty payments at source may create incentives for taxpayers to attempt to change their behaviour to inflate the royalty payments received from a particular jurisdiction. Say, for example, a Canadian company incorporates a subsidiary in India. Payments from the subsidiary to the parent company in the form of interest or dividends may be subject to withholding tax; however, assuming no source-based taxation of royalty payments, a payment for the use of a patent may not. Presumably this approach would create an incentive for the Canadian parent to transfer intangible property to the subsidiary, to value the property at the highest value possible, and to receive royalty payments for its use – thereby managing to remove profits from the subsidiary and back to Canada without the imposition of withholding tax. This kind of strategy would become more appealing if the royalty payment were not subject to tax in Canada, say, because the Canadian parent was non-taxable or had losses.

C. Avoiding Exemptions from Royalty Taxation

The royalties article may exclude particular types of payments from taxation in the country of source altogether. The purpose for this kind of exemption might be two-fold. First, it might be designed to move the taxation of royalty payments slowly closer to the OECD model convention, which removes royalties from the scope of source-based taxation altogether, or it might be justified on the grounds that some types of royalty payments have particularly strong economic links to residence jurisdictions. For example, it has been argued that copyright royalties should only be taxed in the country of source because they are the result of cultural efforts, and the residence country has a stronger claim to the revenue derived from those efforts than the country of source. This argument might be understood two ways. First, it might be understood as an argument about underlying principles – royalty income arising from the use of cultural property has a markedly stronger economic nexus to the residence country than other kinds of property and therefore is justifiably only taxed in the country of residence. While this may be true, as noted above, it does not justify non-taxation at source, although it may justify a higher weighting to residence than source taxation, at least in theory. However, distinguishing cultural property giving rise to royalties from other forms of property may create additional administrative and conceptual difficulties that are not outweighed by the advantages of preserving a higher tax allocation to the residence country from a principled perspective. The UN commentary also suggests that taxation at source of the royalties derived from the use of films, in particular, may be justified as a

---

30 One could reply that the transfer pricing rules may operate to reduce the payment to one that arm’s length parties would agree to; however, the transfer pricing rules are notoriously hard to apply to intangible transfers.

31 See commentary on the UN model convention, paragraphs 10 and 11.
proxy for the taxation of salaries of actors and other participants in the film only in the
country of residence.\footnote{See UN commentary at para 10.} Second, the claim that royalties from cultural property should not
be subject to source taxation might be understood as justified because non-taxation at
source would provide an incentive to produce and disseminate cultural property. To the
extent that the claim that royalties from cultural property should be exempt from source-
taxation is properly characterized as an incentive, that argument is addressed below.

Second, some royalties may be specifically defined as excluded from the royalties
provision (and therefore excluded from source taxation) to provide incentives for
particular transfers. In order for the effect of the exemption to actually provide those
incentives, the type of royalty payment would also need to be exempt from tax
domestically – otherwise the exemption would simply function as a revenue transfer from
the source country to the residence country, and not as an incentive. However, this would
only be the case where either the residence country exempted the particular type of
royalty payment from tax (or taxed it at a comparatively lower rate), or where the
withholding tax imposed would exceed the foreign tax credit in the residence country.
Otherwise, as noted in the introduction of this chapter, a taxpayer should be neutral (from
a tax perspective) between locating property giving rise to royalties in the residence
versus the source country. Instead, that property should simply be located where it makes
most sense in the absence of the tax imposed. Of course, where a residence country
wishes to provide an incentive for the production of particular types of property giving
rise to royalty payments it can always choose to do so through other means, for example,
by implementing a spending program to assist in the production of property that gives
rise to particular types of royalty payments.

Developed countries have also justified the non-imposition of withholding tax on royalty
payments on incentive grounds. For example, countries have argued that avoiding
withholding taxes on royalties better fosters investment. Presumably this argument is
predicated on the assumption that the administrative burden of reporting and collecting
tax in more than one jurisdiction in some cases will be sufficient to dissuade the investor
from transferring the property to the foreign (source) jurisdiction. [continue]

III. Comparing the Tax Treatment of Royalty Payments in Canada and Australia’s
Tax Treaties

This part of the chapter compares the royalties provisions in Canada’s 88 tax treaties to
the 42 tax treaties negotiated by Australian authorities. The comparison focuses on the
various aspects of the royalties article discussed in Part II. Therefore, the following
sections address the taxation (and rate) of withholding at source, and the exclusions of
some types of payments from the royalties withholding tax. While both Canada and
Australia’s tax treaties permit the taxation of royalty payments at source, Canada’s tax
treaties generally impose a higher rate of withholding, and Canada is more likely to
permit non-reciprocal withholding tax rates where the treaty partner is a developing country. Canada is also significantly more likely to have exempted particular types of royalty payments from withholding tax, or to have reduced the withholding tax rate on particular types of payments.

For each of the charts summarizing the results of the comparison between Canada and Australia provided below, the tax treaty partners of each country have been classified as developing, high developing, or low developing. A list of which countries fit within each classification is provided at Appendix A. Although a number of organizations, including the United Nations, the World Bank, and the International Monetary Fund, have classified countries as developing or developed, for the purposes of this chapter, I used a relatively simple approach. If the country had GDP per capita in excess of $17,000 per year, it was classified as developed; if the GDP per capita was between $5,000 and $16,999 it was classified as high developing; and, if it was lower than $4,999 the country was classified as low developing. This is an admittedly rough method of classifying jurisdictions; however, it seemed the most easily justified approach given that my argument is that Canada (and other developed countries) should provide for a more just allocation of tax revenues by permitting developing countries to take a larger slice of those revenues. Gross domestic product per capita seemed like a sensible measure of economic impoverishment.

Chart 1 sets out how many of the tax treaties entered into by each country have been entered into with developed, high developing, or low developing countries. Canada has entered into proportionately more tax treaties with both high developing and low developing countries than Australia.

**Chart 1: Number of Tax Treaties in Force in Canada and Australia, by Development Classification**
In addition, Canada has been entering into tax treaties with developing countries for longer than Australia. For example, Canada entered into its first tax treaty with a high developing country, the Dominican Republic, in 1976, and with a low developing country, Pakistan, in 1974. In contrast, Australia entered into its first tax treaty with a high developing country, South Korea, in 1984, and with a low developing country, the Philippines, in 1980.

A. Taxation of Royalty Payments at Source

In each of their tax treaties, both Canada and Australia permit withholding tax on royalty payments at source, following the UN model convention; however, Canada generally has entered into treaties that have permitted a higher withholding tax rate than Australia. Chart 2 provides a summary of the withholding tax rates agreed to by Canada in its tax treaty negotiations. It shows that the more developing the treaty partner, the more likely Canada was to have permitted a higher withholding tax. The same trend holds true for the Australian treaties, summarized in Chart 3.

Chart 2: Withholding Tax Rate on Royalty Payments in Canada’s Tax Treaties
One of the most progressive steps a country can take in allocating tax revenues to developing countries is to provide for non-reciprocal withholding tax rates. In this case, the developed country taxes royalty payments with a geographic source in their jurisdiction at a lower withholding tax rates than royalty payments with a geographic source in the developing country. This allows the developing country to claim more of
the tax revenues derived from royalties earned in the developed country by its’ residents, and allows it to claim a larger share of the tax revenues earned by developed country residents where those tax revenues have a source in their jurisdiction. Chart 4, below, shows the number of treaties where Canada and Australia have granted non-reciprocal withholding tax rates to their treaty partners.

**Chart 4: Non-Reciprocal Withholding Tax Rates, Canada and Australia’s Positions Compared**

Chart 4 clearly reveals that Canada has been significantly more willing than Australia to grant non-reciprocal withholding tax rates. Quite understandably, neither Australia nor Canada has granted non-reciprocal rates to another developed country. Canada, however, has granted a non-reciprocal withholding tax rate on dividends and interest to a high developing country, Thailand, but has not granted any non-reciprocal withholding tax rates on royalties to high developing countries. Both Canada and Australia have been more likely to allow non-reciprocal rates in the case of dividend payments than for interest and royalties. Although Australia has granted non-reciprocal rates for dividend payments in two of its treaties, it has not granted that benefit to any country for interest or royalty payments, while Canada has granted nine countries non-reciprocal rates on dividends, four countries non-reciprocal rates on interest, and three countries non-reciprocal rates on royalties.

33 Where a tax treaty did not impose a withholding tax on dividends paid from a treaty partner of Canada or Australia at all, because dividends were not subject to tax domestically in that jurisdiction, the non-imposition of a withholding tax on dividends was not classified as a “non-reciprocal” rate.

34 Papua New Guinea and Vietnam [proper cites].

35 Pakistan, Jamaica, Zimbabwe, Philippines, Cameroon, Egypt, Senegal, Papua New Guinea, Ivory Coast [cite properly.]
B. Exemptions for Particular Types of Royalty Payments

Canada and Australia differ significantly in their willingness to exempt particular types of royalty payments from tax, and their willingness to provide a lower withholding tax rate for particular types of royalty payments. As Chart 5 reveals, Australia provides no exemptions from withholding for particular types of royalty payments, in contrast to Canada, which provides significant exemptions, particularly for developed countries.

Chart 5: Exemptions from Withholding Tax on Royalties at Source, Canada and Australia’s Positions Compared

Since Australia does not provide any exemptions from withholding, Chart 6 simply shows the kinds of payments that Canada exempts from tax. Chart 6 contains information about only the tax treaties where Canada has granted an exception; in other words, the tax treaties where Canada provides for no exemptions from withholding tax have been removed from the summary.

Chart 6: Types of Royalty Payments that are Exempt from Withholding Tax Under Canada’s Tax Treaties

---

36 Cameroon, Guyana, Senegal, and Pakistan [cite properly.]
37 Pakistan, Cameroon, and Philippines (proper cites)
Canada commonly grants four exemptions from withholding tax on royalties in its tax treaties. The exemptions in the Canada-Germany Tax Agreement provide typical wording for the four exemptions tracked in the Chart 6: withholding tax exemptions for copyright royalties or like payments for literary, dramatic, musical or artistic work; withholding tax exemptions for royalties for the use of, or the right to use, computer software; withholding tax exemptions for the use of or the right to use any patent; and withholding tax exemptions for the use of or the right to use information concerning industrial, commercial or scientific experience. Article 12.3 of the Canada-Germany Tax Agreement states³⁸:

3. Notwithstanding the provisions of paragraph 2, royalties arising in a Contracting State and paid to a resident of the other Contracting State who is the beneficial owner of the royalties shall be taxable only in that other State if they are:

(a) copyright royalties and other like payments in respect of the production or reproduction of any literary, dramatic, musical or artistic work (but not including royalties in respect of motion picture films nor royalties in respect of works on film or videotape or other means of reproduction for use in connection with television broadcasting);

(b) royalties for the use of, or the right to use, computer software or any patent or for information concerning industrial, commercial or scientific experience (but not including any such royalty provided in connection with a rental or franchise agreement).

Almost all of Canada’s tax treaties with an exemption from withholding grant the exemption for copyright royalties related to cultural property. The other three common exemptions are granted in roughly even numbers, but to a lesser extent than the cultural

³⁸ Insert full citation.
exemption. In three cases, Canada has granted unique exemptions (labelled as “E” in Chart 6). Canada’s treaty with Finland, the country with exemptions AE, provides an exemption for films with a strong cultural connection to either Canada or Finland. Canada’s treaty with the United States, the country with the exemptions ABCDE, provides an exemption for “payments with respect to broadcasting as may be agreed for the purposes of this paragraph in an exchange of notes between the Contracting States”. Finally, Canada’s tax treaty with Malaysia, the high developing country with only an “E” exemption, exempts “approved industrial royalties derived from Malaysia by a resident of Canada who is the beneficial owner thereof” from Malaysian tax.

Lastly, there are some types of payments that are not specifically exempt, but that are subjected to a lower withholding tax rate than other types of royalties. Chart 7 shows the types of royalty payments that benefit from comparatively reduced withholding tax rates in Australia and Chart 8 shows the reduced withholding tax rates agreed to by Canada. As has been common throughout, Canada is more likely than Australia to have granted a reduced withholding tax rate on particular types of payments.

**Chart 7: Royalty Payments that Benefit from Reduced Withholding Tax Rates, Australia**

---

39 This chart does not control for the date that Canada entered into its tax treaties – the exemptions for computer software, patents, and scientific knowledge may have become increasingly used after the 1992 and 1993 Federal Budgets, while the exemptions for cultural property became common to grant after the 1977 Budget. Despite this articulated change in policy, however, there does not appear to be an obvious relationship between the date of entry into force of these treaties and the degree to which Canada has been willing to negotiate exemptions in addition to the exemption for cultural property.

40 Can-Finland at 12.4 provides: “Notwithstanding the provisions of paragraph 2, royalties in respect of cultural motion picture films arising in a Contracting State and beneficially owned by a resident of the other Contracting State shall be taxable only in that other State. This provision shall apply only to royalties (a) paid to a resident of Finland in respect of Finnish films which are approved by the relevant authorities and which are exempt from motion picture tax under the *Motion Picture Tax Act* of 26 June 1964, No 366; (b) paid to a resident of Canada in respect of films wholly or principally directed and produced in Canada and which are included in the list of films prepared by the Committee of the Bureau of Film Festivals established under Order-in-Council PC 1968-400 dated February 29, 1968.”

41 Can-US at 12.3(d) [cite].

42 Canada-Malaysia XII.3 [cite].
As Chart 7 shows, Australia has only granted a reduced rate of withholding tax in three of its tax treaties – those with Argentina, India, and Indonesia.

**Chart 8: Royalty Payments that Benefit from Reduced Withholding Tax Rates, Canada**

It is not surprising that Canada grants no reduced rates of withholding for developing countries, since in such a large portion of Canada’s tax treaties with developing countries it has provided an exemption for the same kinds of payments. In the context of reduced
rates of withholding, Canada is roughly equally likely to have granted a reduced rate of withholding for each of the four primarily types of payments listed above. In a few cases, Canada has also granted a different rate of withholding for trade marks (labelled “F” in Chart 8).43 Finally, and in one case, Canada has granted a reduce rate of withholding for a unique type of royalty payment. In the case of the Canada-Argentina treaty, the regular withholding tax rate of 15 per cent is reduced to 3 per cent of the gross amount paid for the use of, or the right to use, news.44

IV. Some Modest Suggestions

[to be inserted]

---

43 The states with a differential rate for trade marks include Argentina, Brazil, and Tunisia. It might be noted that in some case the differential rate is higher than the “normal” rate, for example, in the case of Tunisia.

44 Insert proper cite.
Appendix A: Treaty Partners Type of Country

[To be inserted.]
A Diversity Theory of Charitable Tax Exemption - Beyond Efficiency, Through Critical Race Theory, Toward Diversity*

David A. Brennen**

ABSTRACT

“A Diversity Theory of Charitable Tax Exemption” is an initial attempt at understanding the charitable tax exemption beyond the borders of the law and economics vision of efficiency. Efficiency is the hallmark of many of the existing theories of charitable tax exemption. This is especially true of theories promulgated prior to 1990 by such notable scholars as Boris Bittker and Henry Hansmann. “A Diversity Theory of Charitable Tax Exemption” articulates a rationale for the charitable tax exemption which is based, not on efficiency, but on notions of diversity and creativity. This alternate theory is called “contextual diversity.”

Contextual diversity theory posits that the charitable tax exemption is a means of diversifying the market and, thus, allowing for more creative and wealth producing opportunities. Further, the exemption is subject to contextual constraints that act to limit the scope of charitable activity. Thus, the article explains how the law and economics concept of efficiency fails to fully rationalize many aspects of the charitable tax exemption that are not amenable to an efficiency analysis. Using an alternate theory of law termed “Law and Market Economy Theory,” the article moves beyond efficiency analysis and uses a variety of interpretive frameworks to more fully explain the charitable tax exemption. As an example, the article uses Critical Race Theory as a means for understanding certain aspects of the charitable tax exemption from a different perspective, such as the public policy doctrine. The public policy doctrine prohibits charitable activity which is inconsistent with “established public policy.” Among the implications of contextual diversity is that the public policy doctrine be eliminated and replaced by a rule that explicitly prohibits invidious racial discrimination by charities while still permitting charities to engage in race-based affirmative action.

*©2006 David A. Brennen.
**Ellison C. Palmer Professor of Tax Law, Mercer University School of Law. B.B.A. Florida Atlantic University; J.D., L.L.M., University of Florida College of Law.
# Table of Contents

**Introduction** ................................................................. 1

I. **Intersections of Economic and Critical Theories of Law** ......................................................... 4  
   A. *Law and Market Theory - Sustainable Wealth* ........ 5  
   B. *Critical Race Theory* ................................. 9  
   C. *The Relatedness of Law and Market Economy and Critical Race Theory* .............. 11

II. **A Diversity Theory of Charitable Tax Exemption** ................................................................. 13  
   A. *Diversity as a Value* ................................. 13  
   B. *Diversity in Context* ................................. 15

III. **The Charitable Tax Exemption - More Than Just Tax Relief** ................................................. 19  
   A. *The Exemption* ........................................... 20  
   B. *Benefits of the Exemption* ............................. 22  
   C. *Obligations that Flow from the Exemption* ......... 24  
      1. *Affirmative Obligations* ....................... 25  
         a. “organized” and “operated” .............. 25  
         b. proper (or charitable) purpose ...... 26  
      2. *Negative Obligations* ......................... 31  
         a. politics and lobbying .................. 31  
         b. private inurement/benefit .......... 32  
   D. *Commercial Competition* .............................. 36  
   E. *Private Foundations* ................................. 37

IV. **Efficiency Theories of Charitable Exemption** ................................................................. 41  
   A. *Public Benefit Subsidy Theory* .................... 41  
   B. *Base-Defining Theory* .............................. 44  
   C. *Capital Formation Subsidy Theory* .............. 50

V. **Potential Implications for the Structure of Tax Exempt Charity Law** ........................................ 55

**Conclusion** ................................................................. 56
Introduction

What is the normative rationale for the federal income tax exemption for nonprofit charitable corporations? Even though the exemption dates back to 1894, Congress has failed to fully rationalize it. Though scholars and courts have attempted over the years to come up with a coherent rationale for the charitable tax exemption, their attempts are focused almost exclusively on economic efficiency. Thus, the charitable tax exemption is typically framed by noted tax scholars like Boris Bittker, Henry Hansmann and others as an economically efficient means of providing certain goods and services to the public. Rationalizing the charitable tax exemption in economic terms is certainly appealing and deceptively comforting. Indeed, since taxation is facially concerned with money, why not articulate a basis for tax exemption in money terms - hence,

economic efficiency.

But no matter how appealing efficiency might seem, it is axiomatic that law, and more particularly tax law, is about much more than economic efficiency. Tax law is about broader conceptions of justice, fairness, and other aspects of a democratic society that extend beyond economic principles. Likewise, the tax exemption for charities is about much more than money, economics or optimal profit. Instead, the charitable tax exemption is principally about accomplishing a value based mission. That mission may at times be at odds with the notion of a pure profit motive that dominates the private market narrative. While efficiency analysis may be relevant to some aspects of a charitable mission, there are many other non-economic aspects of “mission” that extend beyond economics. Thus, this Article does not dispute that traditional efficiency analysis adds to our normative understanding of tax-exempt charity law. Economic analysis may, for example, aid in understanding the exemption’s economic impact. However, efficiency alone simply does not fully

\[4\] For example, even though the federal estate tax burden is born almost exclusively by wealthy taxpayers, its temporary repeal has been promoted politically as, alternatively, an end to the death tax and an end to a tax on small family farms. See, e.g., Sarah Waldeck, An Appeal to Charity: Using Philanthropy to Revitalize the Estate Tax, 24 Va. Tax Rev. 667, 695-96 (2005) (referring to the estate tax as the “death tax” and noting that aside from its aim of preventing high concentrations of wealth, the estate tax has many emotional components). See also, David A. Brennen, Race and Equality Across the Law School Curriculum: The Law of Tax Exemption, 54 J. Leg. Educ. 336 (2004) (explaining how incorporation of concepts of race into legal disciplines such as tax law might enrich study of these areas).

\[5\] In emphasizing the distinctive characteristics of non-profits, including tax-exempt charities, from for-profit organizations, Professor Robin Paul Malloy has explained:

One can reasonably ask why it is that we provide special local, state, and Federal tax benefits to non-profits as opposed to private or for-profit organizations... [M]any believe it is important to do so because non-profits tend to supply public goods, the provision of which is usually under-provided in the private market. Furthermore, non-profits generally seek to promote values that are difficult to measure in economic terms. The non-profit framework, therefore, raises some important cultural-interpretive issues. The focus on values, and the rejection of a pure profit motive, are two very important points of diversion from the conventionalized norms expressed in our private market narratives.

See ROBIN PAUL MALLOY, LAW IN A MARKET CONTEXT: AN INTRODUCTION TO MARKET CONCEPTS IN LEGAL REASONING 214 (Cambridge University Press 2004) (emphasis in original) [hereinafter MALLOY, LAW IN A MARKET CONTEXT].
explain the varied and rich non-economic aspects of the charitable tax exemption. Accordingly, this Article offers an alternative framing of the charitable tax exemption that serves as an alternative to the longstanding economic theories of scholars like Boris Bittker, Henry Hansmann and others. The Article demonstrates that a principal normative justification for the exemption, in addition to economic efficiency, is diversity - what this Article calls “contextual diversity.”

Part I of the Article presents Robin Paul Malloy's Law and Market Economy Theory ("LMT") as an example of the basis for a normative explanation of the charitable tax exemption. LMT addresses the relationship among law, markets and culture. Thus, using LMT, this part demonstrates how traditional law and economic analysis, premised on self-interest and wealth maximization, simply does not capture the essence of the many values that impact the marketplace and the market exchange process. Instead, LMT approaches legal analysis in a broader market context and is premised on the need to promote a process of sustainable wealth formation as opposed to maximizing wealth. As an illustration of the difference between traditional law and economics and LMT, Part I addresses the way in which LMT is compatible with Critical Race Theory in discerning normative rationales for law that enable richer reflections of justice, fairness and equality that are not necessarily connected to positive economics or efficiency.

Part II of the Article presents a theory of the charitable tax exemption that is line with LMT and based on the value of diversity. "Contextual diversity," as explained herein, references alternative understandings of LMT to build on insights of traditional economic analysis and present an alternative value-based rationalization for the charitable tax exemption. Part II concludes by explaining that using contextual diversity as a principle value based rationalization for the charitable tax exemption not only captures the essence of the exemption, but it also provides direction for potentially reforming and re-inventing various aspects of tax-exempt charity law.

Part III of the Article provides a detailed outline of the charitable tax exemption - explaining both the affirmative and negative aspects of the exemption. The purpose of this detailed outline is to position the exemption as not simply a financial “free pass” on the obligation to pay income tax. Rather, the charitable tax exemption is the gateway to an alternative way of operating an enterprise, with many burdens and benefits flowing therefrom that often have nothing to do with economics or efficiency. Instead, these non-economic aspects of the charitable tax exemption often concern
the many non-economic aspects of justice, fairness, equality, political authority and other basic normative principles. Tax-exempt charities are not like private for-profit firms that measure success by bottom-line profits; nor are they like public entities that measure success by voter support or re-election. Tax-exempt charities measure success by how well their missions are accomplished. Measuring success in this way is inherently “more of a normative judgement than it is for a private entity” or a public entity “because there are fewer external indicators available” for determining a charity’s success.⁶

Part IV builds on Part I by demonstrating how the predominant traditional theories of charitable tax exemption just do not capture the full essence of the normative aspects of the exemption. Part IV analyzes the traditional theories promulgated prior to 1990 which rely principally on economic efficiency as a guidepost for discerning an appropriate rationale. While these traditional theories have a sexy “law and economics” appeal, the Article demonstrates how they fail to fully explain many non-economic aspects of the charitable tax exemption.⁷ Finally, part V briefly identifies some of the implications of contextual diversity theory on the structure of tax exempt charity law.

I. INTERSECTIONS OF ECONOMIC AND CRITICAL THEORIES OF LAW

⁶See Id at 220-21.
⁷This Article focuses primarily on the pre-1990 theories of charitable tax exemption because they tend to rely almost exclusively on notions of economic efficiency. In future research, I intend to examine theories promulgated since 1990 that challenge the pre-1990 traditional theories. These more recent theories of exemption challenge the traditional theories by focusing less on economic efficiency and more on non-economic values. They draw on notions of philosophy, political theory and moral values as a means for complementing the traditional economics focus of the pre-1990 theories. Thus, these recent theories explain the charitable tax exemption from the perspective of the selflessness (as opposed to self-interest) of the donor, see e.g., Atkinson, *Altruism*, supra note 3; Atkinson, *Theories*, supra note 3, and the alternate political authority offered by the charitable form, see e.g., Brody, *supra* note 3. These other theories of the charitable tax exemption come closer to articulating the type of contextual diversity contemplated in this Article. For they implicitly recognize that much of the justification for the charitable tax exemption is not about economic efficiency. Many aspects of the exemption concern notions of justice, fairness and opportunity - concepts that are not necessarily translatable into economic or efficiency terminology. However, I intend to develop, in later research, the idea that even these theories do not fully account for the many complexities inherent in the charitable tax exemption.
LMT and its interplay with Critical Race Theory (CRT) is just one example of how different theoretical approaches to law can come together to partially expose important elements of a particular substantive area of law. Sub-part A outlines LMT - a theory of law that is based on economic principles, but which extends beyond traditional economic analysis to incorporate other humanistic values into an understanding of law and social interaction. LMT posits that law is more fully understood when economic analysis is combined with other approaches to legal analysis to demonstrate the dynamic and ever changing aspects of human interaction. As one example of how other theories of law might complement economic understandings of the law, sub-part B demonstrates how CRT might be used as a means of understanding aspects of law that are not fully understood by economics alone. Overall, this section demonstrates how traditional understandings about economic analysis of law could be diversified by accommodating other approaches to law, such as CRT. Such accommodation could lead to better, more diverse, understandings of the structure of tax-exempt charity law and potentially improve future development of this burgeoning area of law.

A. Law and Market Theory - Sustainable Wealth

Traditional economic analysis of law is premised on the idea that efficiency leads to wealth maximization. Richard Posner, in his book *Economic Analysis of Law*, writes that: “economics is the science of rational choice in a world in which resources are limited in relation to human wants.” This means that, given the scarcity of resources that people desire, economic analysis attempts to predict how best to allocate these resources. If these resources will be allocated best without legal intervention, then law and economics would suggest that no law be created. On the other hand, if law is needed to ensure proper resource allocation, law and economics would so indicate. Thus, according to Posner, law and economics has both positive and normative aspects. On the positive side, economic analysis can be used to describe legal rules and results that form the foundation of common law. Normatively, economic analysis asserts

---


that law makers and judges should opt for legal rules that advance efficient allocation of legal resources. The job of economics, then, is to examine the results of assuming that humans are rational maximizers and, hence, naturally pursuers of self-interest. Economic analysis suggests that in pursuing self-interest, individuals will necessarily advance the public interest. Thus, the goal of economics is efficiency, or in other words, the allocation of resources in such a way as to maximize value - where value is how much someone is willing to pay for something.10

For a number of reasons, economic analysis is an inappropriate means, by itself, for making decisions about the proper structure of law. Indeed, law is about social interaction. Thus, law is necessarily about justice, fairness, equality and other aspects of social existence. Economic analysis often fails to capture all of these aspects of law. The primary reason for this failure of economic analysis is its attempt to make purely scientific that which cannot be explained by science alone. That is, law and economics as an approach to legal decision making seeks to make law appear more, rather than less, scientific.11 Thus, traditional law and economics avoids resort to humanities and other disciplines that provide alternative explanations of human interaction and social constructs. This is not to say that economic analysis is not useful for legal reasoning and decision making. To the contrary, economic analysis has a definite role to play, but that role must accommodate for the many other aspects of human existence aside from those that are amenable to precise scientific understanding.

Specifically, one of the major failings of traditional economic analysis is the absence of accounting for decision making that is not “rational” in an economic sense, but yet, is fully justifiable on other grounds. People are not always motivated to act by self-interest alone. Indeed, people often act for purely non-selfish reasons - such as out of concern for the well-being of others, out of habit, in response to tradition, or out of regard for their subordination to the power of others. These various reasons for human action require reference to more than economics to be properly understood. Thus, to structure laws based primarily on the sole assumption of people as self-interested actors would not necessarily respond to real societal

\[11\text{Robin Paul Malloy, Framing the Market: Representations of Meaning and Value in Law, Markets, and Culture, 51 Buff. L. Rev. 1, 17 (2003) [hereinafter Malloy, Framing the Market].}\]
needs. Charities operate, not out of private self-interest, but instead out of a public benefit, or mission, interest. This is distinct from for-profit corporations, which operate so as to maximize economic profits. Therefore, attempts to articulate a rationale for the charitable tax exemption by resort to economic analysis alone necessarily miss this important distinguishing aspect of charitable operations.

In his "law and market economy" theory, Professor Robin Paul Malloy echoes a similar theme as that advanced by Nobel prize winning economist Amartya Sen in explaining that efficiency and wealth maximization are inadequate measures for assessing social well-being. Malloy explains the market as a place of meaning and value formation in which real value emerges from the continuous process of exchange and interaction. He argues that the process of sustainable wealth formation is difficult to measure and understand in traditional efficiency terms. He demonstrates that the process of sustainable wealth formation relies on creativity, and on the dynamic nature of inclusive, diverse, and extensive networks and patterns of exchange. Value emerges from the dynamic interaction and play between people and their ideas. And, wealth formation is facilitated by connecting diverse individuals and communities so that valuable information fragments and experiences enter into the broader marketplace. Discrimination, exclusion, and inaccessibility hinder the sustainability of the wealth formation process.

In this regard, Malloy challenges three fundamental aspects of traditional law and economics concerning the primary tension in law and economics, the primary means of wealth formation in the market, and the nature of market choice. First, Malloy explains that the primary tension in LMT is between the concepts of efficiency and creativity, not between efficiency and social responsibility as alleged by traditional law and economics. He says that there is no need to understand the counterpart of social responsibility as efficiency because a properly functioning market incorporates an ethic of social

---

12 See Malloy, Law in a Market Context, supra note 5, at 50-52.
14 See Id. at 124.
15 Id. at 2-4. More specifically, law and economics alleges that the primary tension in the market is between efficiency and social responsibility; that efficiency is the primary means of wealth formation in the market; and that market choice is rational, objective and the scientific result of cost and benefit analysis.
responsibility. According to Malloy, efficiency, in law and economic terms, is grounded in the static notions of habit, convention and continuity. Efficiency is reactive and is grounded in making the most of current understandings of the market. Creativity, on the other hand, is much more dynamic. According to Malloy, creativity is grounded in notions of potentiality, discontinuity and indeterminancy. Thus, creativity is necessarily proactive and ever-evolving.16

Second, drawing on the tension between efficiency and creativity, Malloy argues that creativity in the market is the primary means of wealth formation, not efficiency. While efficiency is certainly relevant and has an important role to play in market theory, efficiency just cannot account for the process of creativity. Efficiency requires ideal environments based on habit informed conventions and pre-established relational choices. Creativity, on the other hand, relies on habit deforming and transforming exchange relationships that permit the discovery of something new and different.17 Creativity is by nature indeterminate, habit-breaking, and convention-challenging. Since it cannot be observed directly, creativity can only be examined by looking at the contextual communities which foster it. As Professor Malloy explains: “One must identify the types of communities which, by ethics and social values, tend to foster diversity, experimentation, and unconventional networks and patterns of exchange.”18 These are communities that embrace inclusion and diversity, and are ones that think about the market process in terms that are broader than economic efficiency.

Third, Malloy explains that market choice involves a process of interpretation - it is not a rational or objective fact which can be determined scientifically by cost and benefit analysis, unimpacted by social influences. Rather, market choice is the result of one’s interpretation of market incentives and disincentives as informed by personal experiences rather than by abstract notions of objectivity and rationality.19 A person’s cultural biases, whether relating to race, sex or sexual orientation, for example, have a decided impact on how he/she defines, interprets and weighs the costs and benefits of a particular action. This distinction between rational choice and interpretation is a key aspect of Professor Malloy’s divergence from traditional law and economics. As Professor Malloy explains:

---

16Id. at 2-3.
17See Malloy, Law and Market Economy, supra note 13, at 78-90.
18Id. at 3 (emphasis added).
19Id. at 3-4.
This distinction is important because the process of interpretation is community based, and because it indicates that even though exchange takes place as a continuous part of a dynamic system, our understanding of the exchange process is shaped by the interpretive “lens” or “screen” through which we view it. Furthermore, this lens or screen, as an indexical reference in semiotics, is grounded in a system of values informed by experience rather than by purely objective and rational choice.\footnote{This is a reference to footnote 10 of the source text.}

In essence, the reliance in law and economics on methodological individualism is flawed because it extracts the individual from the cultural-interpretive community in which she is embedded. At the same time it disconnects the individual from history and context, thus ignoring the interplay between an individual's experiences and her cultural-interpretive framework for understanding.\footnote{This is a reference to footnote 21 of the source text.}

\subsection*{B. Critical Race Theory}

Instead of simply relying on economic principles to articulate a coherent rationale for the charitable tax exemption, traditional law and economic analysis should (and could) - as urged by LMT - be combined with other legal approaches to law in order to add an additional \textit{screen} or \textit{lens} through which to rationalize charitable tax exemption law. One example of a legal approach to law that might provide an additional lens, in addition to economic analysis, through which to filter and understand the charitable tax exemption is CRT.

CRT developed in the latter part of the twentieth century as a response by progressive scholars of color to critical legal studies, which in turn was a response to Legal Realism.\footnote{This is a reference to footnote 22 of the source text.} Legal Realists...
criticized the rule of law as overly formalistic and promoting, under the guise of neutrality and objectivity, a system of laws actually driven by policy, economics and politics. These revelations ultimately lead Realists to pronounce the law as indeterminate by nature. Critical legal scholars took the Realists agenda a step further by deconstructing and deligitimizing the law - describing it as not only political, but also ideological and hegemonic. As time passed, some scholars of color became disenchanted with critical legal studies because of its failure to acknowledge White supremacy and otherwise meaningfully critique racial power and racial hegemony. These progressive scholars of colors developed their own race-conscious approach to legal analysis, seeking to expose the legal and social construction of race itself with the goal of eradicating racial subordination.

CRT is characterized primarily by its opposition to three mainstream beliefs about racism: 1) that color-blindness will eliminate racism; 2) that racism is an individual act, not a systemic problem; and 3) that problems of racism can be addressed without dealing with other forms of societal injustice such as sexism, homophobia, or economic exploitation. One of the central concepts in CRT is that race is a social construct - an idea. In other words, the concept of race occurs, not naturally, but by invention in society. Understanding race as a social construct helps to explain both racial existence and racial hierarchy. Another important concept in CRT is its rejection of racial essentialism. Essentialism refers to the assumption that a particular race has a certain essence. CRT rejects this essentialist approach to race. Instead, for CRT, race is contextual. As Kimberle Crenshaw has explained, CRT is committed to the idea that racial identities are intersectional and that racial minorities’ vulnerability to discrimination is a function of specific intersectional identities.
C. The Relatedness of Law and Market Economy and Critical Race Theory

Malloy’s law and market economy theory shares many attributes with CRT on the subject of challenging traditional law and economic analysis. 30 Both law and market economy theory and CRT reject the primacy of efficiency as espoused in traditional economic analysis. 31 Professor Malloy’s challenge to efficiency is from a market perspective, while critical race theorists challenge efficiency from an equality perspective. 32 Professor Malloy explains that creativity, not efficiency, is the primary means of wealth creation in the market. Similarly, critical legal scholars explain that racism is not just a problem of individual choice, but instead is the result of a systematic condition of racial subordination. Thus, both law and market economy theory and CRT reject the notion that the focus of proper legal structure should be on the calculus of individual choice. Instead, the focus should be on the diversity of societal structures that create the circumstances leading to the choice.

According to Malloy, market analysis is primarily concerned with the human practice of exchange - about who initiates exchange, who is excluded from exchange, what is subject to exchange, and on what terms exchange occurs. 33 Furthermore, he argues that it is important to examine the way in which people experience the market exchange process - noting that experience varies by a number of characteristics such as race, gender, age, and education level, among others. 34 These contextual and experienced-based considerations of the market are consistent with similar considerations and values expressed in CRT. 35

In addition to challenging traditional law and economics as focused too much on efficiency, law and market economy theory and CRT both acknowledge the overt political nature of market actors - affirmatively acknowledging the societal/contextual influences on law

30 See Houh, supra note 22, at 1063-66.
31 See id.
32 See id.
33 See Malloy, Law in a Market Context, supra note 5, at 1-25; Malloy, Law and Market Economy, supra note 13, at 57-70.
34 See Malloy, Law in a Market Context, supra note 5, 1-25.
35 See Malloy, Law in a Market Context, supra note 5, at 6 as an example of these connections.
makers, judges and those subject to law. 36 Traditional law and economics attempts to conceal many of these contextual influences by projecting an heir of neutrality or objectivity through use of the science of economics. However, Professor Malloy unmasks this charade by explaining how law and economics scholars misunderstand both the impact of racial subordination on market exchange processes and the indeterminate nature of the market exchange process. 37 Despite the similarities between CRT and traditional law and economics theory with regard to both being politicized - one overtly so and other not so overtly, many still view law and economics as more objective than CRT. Nevertheless, it is clear that CRT has a role to play in understanding law in a social context.

More specifically, because critical race theorists offer a view of social context that is just as legitimate as that offered by traditional law and economics, it seems perfectly reasonable to examine the charitable tax exemption through the dual lenses of CRT and traditional law and economic analysis. Though tax laws by nature appear readily amendable to economic analysis, this conclusion fails to take account of the uniqueness of tax laws concerning charities. Unlike many tax laws that deal with raising government revenue, tax laws imposed on charities do not have this overt revenue raising goal. Instead, the overarching goal of tax laws imposed on charities is to enhance the mission of charities. In fact, it is only when a charity diverts from its mission that its tax exemption is jeopardized. 38 As one examines the complex nature of the charitable tax exemption, one must ask, “Why are charitable corporations not required to pay income taxes?” The exemption has traditionally been rationalized on efficiency grounds. However, efficiency cannot be used to rationalize all aspects of the charitable tax exemption. What about the other conceptions of justice and fairness that are not reflected in the calculus of economic efficiency? Thus, LMT offers an analytical approach to the charitable tax exemption that helps us to better rationalize it.

36 See Houh, supra note 22, at 1063-66.
37 See id.
38 Cite examples: IRS’s challenge to NAACP’s tax exempt status because of its leader’s speech against president Bush; Litigation involving Branch Ministries tax exempt status instigated by Branch Ministry’s ad in a national newspaper against Bill Clinton; controversy over CBN’s tax exempt status stemming from its commercially successful broadcast network.
In support of this proposition - that LMT permits a better rationalization of the charitable tax exemption, the remaining parts of this Article do four things. First, part II proposes a new theory of the charitable tax exemption premised on LMT concepts. Second, part III explains various aspects of the charitable tax exemption from the perspective of this new theory. Third, part IV compares and contrasts this new theory with many of the pre-1990 theories of the charitable tax exemption that are based almost exclusively on principles of traditional economic efficiency. Finally, part V briefly identifies some of the implications of this new theory on the structure of tax exempt charity law.

II. A DIVERSITY THEORY OF CHARITABLE TAX EXEMPTION

A. Diversity as a Value

Diversity has long been recognized as globally beneficial. In nature, for example, diversity is seen as a key component for plant and animal species survival. In financial investments, diversity is seen as the principal means of reducing risk of capital loss and ensuring sustainable growth over the long run. Importantly, in education, society values diversity as imperative if one is to have a quality and well-rounded educational experience.\(^3^9\) This Article asserts that diversity is also the driving force behind the charitable tax exemption. The diversity made possible by the charitable tax exemption breeds creativity, ingenuity and other things that stimulate society and, in turn, market growth and development. The charitable tax exemption contributes to diversity by offering alternative means of accomplishing societal objectives in a market context. Thus, charities offer alternatives to for-profit corporations and to government.\(^4^0\)

Charities are just one of the many types of corporate entities


\(^{4^0}\)Professor Jill Horowitz has reached similar conclusions regarding the value to society of the alternative forms of hospitals, including for-profit, not-for-profit and government hospitals. See Jill Horowitz, Why We Need the Nonprofit Sector: The Behavior, Law and Ethics of Not-for-Profit Hospitals, 50 UCLA L. Rev. 1345 (2003) [hereinafter Horowitz].
that operate in the United States. A key distinction between nonprofit charitable corporations and for-profit corporations is the lack of shareholders in the former. This means that while for-profit corporations are generally presumed to be profit-motivated institutions, nonprofit charities carry no such presumption. Instead, nonprofit charities are mission driven and are legally prohibited by the nondistribution constraint from extracting profits from the corporation for private or personal gain. Thus, nonprofit charities provide a viable alternative to the for-profit way of running a corporation - mission focused as opposed to profit focused. This alternative corporate structure - indeed, this diversity in corporate structure - in turn, contributes to the possibility and actuality of diversity in output. For example, this diversity in corporate structure means that we have for-profit corporations performing research concerning potentially very profitable drugs for treatment of HIV and AIDS, while nonprofit charitable corporations search for a far less profitable vaccine for this same disease. Imposing a tax on income incidentally earned from these mission driven activities would only serve to slow accomplishment of the mission.

In addition to being an alternative to the corporate form of organization, charities also offer an alternative to government. Like nonprofit charities, government does not have shareholders and is mission driven. However, one key difference between government and all corporations - both for-profit and nonprofit - is government’s ability to impose a tax (through law) as a means of financing the production of its goods and services. The government’s ability to raise capital through taxation does not mean that government has an endless supply of money, nor does it mean that government can do whatever it wishes to do with this money. Indeed, a key limitation of government activity is the requirement to obtain political support. This means that, unless the public at large is willing to support the

---

41 See infra Part III.
42 See NPR transcript from 2003 describing nonprofit efforts to develop an aids vaccine as compared to for profit efforts to seek potentially more financially profitable drugs for treatment of aids.
43 This is one way in which efficiency analysis as a normative principle works in conjunction with other conceptions of justice to help explain the charitable tax exemption. As more fully developed in part IV, this Article embraces the idea that an income tax on charity income would “cut retained earnings... , and hence would further cripple a group of organizations already capital-constrained.” See Hansmann, supra note 3, at 74.
44 See Brody, supra note 3.
acquisition of particular goods or services with direct outlay of tax monies, those goods or services will not be acquired by government and, hence, will not be supplied to the public by government. Accordingly, to the extent that provision of a particular good or service has some support amongst citizens, yet lacks sufficient political support to garner government backing, it will be up to nonprofit charities to supply the good or service - financing that supply with either tax-exempt profits or tax deductible donations.

B. Diversity in Context

Diversity as a value, however, need not be thought of as an unprincipled process approach to law that is lacking in limits or principle. Diversity as a value should be considered in context. Here, “context” refers to that aspect of law that requires consideration of multiple points of interest - both public interests and private interests. For example, diversity as a value does not mean that charities should be able to advance any conceivable private purpose. Thus, tax-exempt charities, because of the public policy doctrine, cannot engage in invidious racial discrimination against black people. Simply recognizing that racial preferences promote diversity is not enough. LMT teaches us that law must also account for racial preferences in various contexts and from various perspectives - that is, law must contextually mediate public societal interests and private societal interests regarding racial preferences. For instance, a critical race perspective of law would indicate that there is a qualitative and meaningful difference between socially beneficial race-based affirmative action (equalizing opportunities) and the societal harm advanced by racial discrimination (continued racial subordination). Thus, contextual diversity would suggest that even though racial preference in the context of racial subordination is not permissible in tax-exempt charity law, racial preference in the context of affirmative action may be permissible.
February 24, 2006: DRAFT ONLY. DO NOT CITE/REPRODUCE/CIRCULATE WITHOUT PERMISSION

“Context” also explains other aspects of charitable tax exemption law which limit private profit taking and private benefit,\(^{50}\) lobbying and political campaigning,\(^{51}\) competitive business functions\(^{52}\) and private endowments.\(^{53}\) Granted, each of these limitations imposed on charities - including the public policy limitation - appears to impinge on what a charity may do and, thus, seem to hamper certain types of diversity. One must bear in mind, however, that the aim of the charitable tax exemption is to allow for activities that benefit public interests, not private ones. Thus, as explained in this Article, the various limitations imposed on charitable activity are all aimed, in one way or another, at contending with the individual human tendency to seek out authoritative influence to advance private objectives.\(^{54}\) In other words, the purpose of these limitations is to mediate between the private individual tendency for authoritative control and the public interests advanced by the charitable entity. Thus, the diversity value advocated in this

In Breyer’s view, racial preference in the context of affirmative action is not only constitutional, but also consistent with basic democratic principles:

TOTENBERG: On the question of affirmative action in college admissions, Breyer defends the practice as an effort to include all segments of a population in higher education or, as one opinion in the majority put it, 'The court should read the 14th Amendment guarantee to equal protection of the law in light of its purpose, to take people who had once been slaves and make them full members of society.' That view prevailed in the Supreme Court by a 5-to-4 vote, but the dissenter argued that the 14th Amendment was intended to create a color-blind society that would judge people on merit, not skin color.

Justice BREYER: Both of those theories, as a matter of pure logic, are good. So in a situation where an interpretation is so close, I think it's useful to turn back to the basic function of the Constitution; that's where I think the democratic purpose had a role. Members of the armed forces, members of the business community, members of the university community said, 'We need at least a limited degree of affirmative action in order to create a society where people will feel they belong to our institution.'

NPR Transcript, September 29, 2005, (Supreme Court Justice Breyer on "Active Liberty") (All Things Considered).

\(^{50}\)See Part IIIC2b.
\(^{51}\)See Part IIIC2a.
\(^{52}\)See Part IIID.
\(^{53}\)See Part IIIE.
\(^{54}\)Malloy, Framing the Market, supra note 11.
Article is not unbridled diversity; rather, it is “contextual diversity” in the sense that the scope of permissible charitable activity is not without limits.

In addition to understanding that diversity as a value is not unlimited in scope, one must also appreciate the fact that the theory espoused in this Article is just that - a theory of the charitable tax exemption. Like all theories, it is not perfect and is not intended to necessarily reflect every aspect of the charitable tax exemption. Indeed, just as economic analysis admittedly does not explain all that bears on the marketplace, the contextual diversity theory outlined in this Article does not profess to explain every aspect of the charitable tax exemption. However, what contextual diversity does do is explain many aspects of the charitable tax exemption that are not explained, nor are explainable, by traditional economic theories. Contextual diversity also complements, and in many cases extends, more modern theories of the charitable tax exemption, such as Rob Atkinson’s altruism theory and Evelyn Brody’s sovereignty theory. Thus, just as the altruism and the sovereignty theories were not intended to replace Bittker and Rahdert’s base-defining theory or Hansmann’s capital formation theory, neither is contextual diversity intended to replace altruism, sovereignty or any economic theory of charitable tax exemption.

In the end, contextual diversity offers a normative rationalization of the charitable tax exemption that can assist - in conjunction with economic analysis - in better outlining the contours of charitable exemption law. Consider again the example of tax-exempt law’s public policy doctrine and its limitation of charitable activity to activity that is not inconsistent with established public policy. Established public policy is often conceived of in application as federal governmental policy or majoritarian compliance. Contextual diversity would suggest that the adoption of the public policy doctrine was inappropriate in the sense that it was overkill. Indeed, “established public policy” - whatever it means - does not define the bounds of charity. Instead, LMT suggests that the scope of

---

charitable activity is varied, diverse, dynamic, and transformative. Charitable activity may be consistent with, have nothing to do with, or, most importantly, be completely contrary to “established public policy” as presently conceived.

A Piercean semiotic interpretation of the legal idea (or, semiotically speaking, sign) “charitable” illustrates that its meaning has variance across cultural-interpretive communities. In this regard the designation of an activity as "charitable" can be understood as an interpretation of, and a representation of, particular underlying values. In some contexts, for example, "charitable" stands as a representation for fulfilling a public purpose with respect to others who are truly in need, with no pejorative connotation. In other contexts the term “charitable” denotes action taken in support of subordinated people and functions as a sign of one's nobility in dealing with lesser or inferior beings. As a sign of public policy, "charitable" activities may take on multiple nuanced meanings and may function, depending upon the context, as an interpretation of underlying socio-legal values supporting invidious racial discrimination. We cannot determine the appropriate contextual meaning of "charitable" by reference only to positive economics and its emphasis on efficiency analysis. We need a more textured and nuanced approach to exchange relationships in a market context.

For example, understanding the problems of permissible and impermissible racial preferences must be informed by the contextual positioning of the parties involved. Markets are not fully objective and neutral avenues of exchange, they are the product of human

---

57See, e.g., Malloy, Framing the Market, supra note 11, at 103.

…… [P]eople positioned in alternative interpretive communities use different interpretive frames and references. Thus, different people understand the world in different and sometimes conflicting ways. Therefore, we must be aware of a variety of cultural-interpretive perspectives as they influence the direction of law and social policy. Consequently, we must understand the relationship between law, markets and culture, and we must realize that by shifting interpretive perspectives we can alter authoritative influence over the interpretation process.

Id.

58Similarly, Professor Houh concludes that, for contract law, “[c]ontractual good faith may mean one thing to a critical race scholar, another to a conventional law and economics scholar, and another to a law and economics scholar.” See Houh, supra note 22, at 1051.

59See Malloy, Law in a Market Context, supra note 5, at 6-16, 104-10 (explaining framing, referencing, and representing).
practice and culturally informed values. Thus, when faced with the issue of the permissibility of invidious racial discrimination by tax-exempt charities, a careful consideration of the context of this type of racial preference reveals that mere racial preference was not the problem in *Bob Jones University v. United States*. The problem, as Critical Race Theory teaches us, was a problem of unjustified inequality - the continued racial subordination of blacks long after the end of legalized slavery. Accordingly, prohibiting racial subordination (the underlying problem), instead of prohibiting acts that are contrary to "established public policy," will better advance the goals of contextual diversity.

With this understanding of the charitable tax exemption as a backdrop, the Article now begins to show how “contextual diversity” implicitly permeates various aspects of the exemption.

### III. THE CHARITABLE TAX EXEMPTION - MORE THAN JUST TAX RELIEF

The charitable tax exemption consists of many varied and complex components. This section of the Article describes these components in order to demonstrate that the charitable tax exemption is more than just relief of a financial obligation to pay a federal tax on income. Indeed, many charities arguably do not have income—or, if they do, they seek tax-exempt status primarily for reasons that have nothing to do with the federal income tax exemption. Thus, the purpose of this section is to show the important and vital role the charitable tax exemption plays in daily life from the standpoint of justice, fairness, equality and other non-economic values. This section also articulates that, along with the general income tax exemption, tax-exempt charities receive many other benefits and must bear many burdens. In short, the charitable tax exemption is about

---

60In recommending the abolition of the public policy limitation in favor of an explicit rule prohibiting invidious racial discrimination by tax-exempt charities, this Article draws on the teachings of Critical Race Theory regarding legal rules and principles that are “originally” prompted by racial discord but that avoid explicitly mentioning race in the formally adopted legal rule or doctrine. See e.g., Jerome Culp, *Toward a Black Legal Scholarship*, 1991 Duke L.J. 39, 67-77 (1991) (“This liberal view of the Constitution and race, that race is better left unexplored, prevailed in much of the constitutional drafting.”).


much more than mere tax relief; it is about how private market actors benefit from having corporations that are profit-focused in addition to having corporations that are mission-focused.\(^63\)

\[A. \quad \textit{The Exemption} \]

At its most basic level, the charitable income tax exemption that is the subject of this Article is the statutory relief from the obligation of certain nonprofit corporations to pay tax on annual income.\(^64\) This relief is granted pursuant to federal law and is given automatically to those corporations that apply for and are granted tax-exempt charitable status.\(^65\) A variety of corporations other than charities are entitled to income tax exemption, including social welfare organizations,\(^66\) labor organizations,\(^67\) business leagues\(^68\) and social clubs,\(^69\) just to name a few.\(^70\) However, charitable nonprofits, unlike practically all other tax-exempt nonprofits, are special in a variety of ways. For instance, with few exceptions, charities are the only tax-exempt nonprofits that are also eligible to receive charitable donations from the public that entitle the donor to federal income tax benefits.\(^71\) Federal law provides that individuals and corporations that donate money or property to charitable corporations may be entitled to receive a tax savings when computing their own tax liability. The potential savings can be quite significant depending on the amount of the donation,\(^72\) the type of property donated,\(^73\) the income of the

\(^{63}\) The essential implications of what it means to be mission-focused are outlined in the section on “proper purpose”. See infra, Part IIIC2.

\(^{64}\) I.R.C. § 501(a) (West 2005).

\(^{65}\) I.R.C. § 501(c)(3) (West 2005).

\(^{66}\) I.R.C. § 501(c)(4) (West 2005).

\(^{67}\) I.R.C. § 501(c)(5) (West 2005).

\(^{68}\) I.R.C. § 501(c)(6) (West 2005).

\(^{69}\) I.R.C. § 501(c)(7) (West 2005).

\(^{70}\) I.R.C. § 501(c) (West 2005) (listing more than 20 different types of tax-exempt organizations.)

\(^{71}\) I.R.C. § 170(a) (West 2005). See I.R.C. § 170(c)(4) (West 2005) for exceptions for special funds in fraternal societies, etc.

\(^{72}\) I.R.C. § 170(a) (West 2005).

\(^{73}\) See I.R.C. § 170(b)(1)(C) (West 2005) (imposing limits on deductibility of certain capital gain property).
This ability to receive tax deductible donations from the public is the key federal tax law distinction between charities and other tax-exempt nonprofits.76

In order to obtain tax-exempt charitable status, a desiring nonprofit corporation must file certain forms with the Internal Revenue Service (“IRS”) seeking such status.77 This is no small task. The forms are very complex and seek a voluminous amount of information concerning proposed organizational structure, proposed activities, financial assets, expected and past revenue streams, proposed policies, and much more.78 Pursuant to federal law, the IRS must evaluate the information provided on these forms to determine if the applicant’s proposed activities and organizational structure are consistent with activities and structures that are deemed “charitable”. Importantly, if the IRS determines that the information provided on the organization’s forms do not adequately comply with the law, then the IRS, subject to later judicial review, has sole discretion as to whether to grant or deny the applicant tax-exempt charitable status.79 Further, even if it grants tax-exempt charitable status based on the applicant’s organizational documents, the IRS can later revoke that status if the applicant does not operate in compliance with federal law.80 Accordingly, even after its initial submission of forms seeking

---

75See I.R.C. § 170(b)(1)(B and D) (West 2005) (imposing limits based on whether the charity is a private foundation or not).
76Although the charitable tax deduction for individuals who contribute to charities is a key aspect of charitable corporations, this Article does not attempt to incorporate the deduction component into an overall rationalization of the exemption. Others who have attempted to draft a theory of tax exemption have similarly omitted discussion of the charitable tax deduction because of the “different issues” raised. See, e.g., Hansmann, supra note 3, at 55.
77I.R.C. § 508(a) (West 2005) requires filing of IRS form 1023.
78IRS Form 1023 requires the applicant to disclose basic identification information. It also requires disclosure of information about the organizational structure, compensation of officers, members of the organization, activities, financial data, public charity status, and user fee information. There are also various schedules that will need to be filed out depending on the type of organization. See IRS Form 1023, available at http://www.irs.gov/pub/irs-pdf/f1023.pdf.
79I.R.C. § 7805(a) (West 2002) delegates to the Treasury department, who further delegates to the IRS, the power to prescribe all needful rules and regulations concerning the enforcement of the internal revenue code.
tax-exempt charitable status, a charity must often also submit annual information reports to the IRS concerning its on-going operations.\textsuperscript{81}

These extensive requirements related to obtaining and maintaining tax-exempt charitable status provide one with just a glimpse of the many aspects of the charitable tax exemption that go far beyond merely being excluded from the requirement to pay federal income tax. Granted, all tax-exempt nonprofits - charitable or not - must file forms seeking tax-exempt status.\textsuperscript{82} However, the forms required for tax-exempt charitable status are much more involved and much more complicated than those required of other tax-exempt nonprofits.\textsuperscript{83} Perhaps the added complexity is due in part to the additional financial impact on tax revenues that accompanies granting tax exemption to an organization that will also be eligible to receive tax deductible donations from the public. But the added complexity could also be related to something else, something that has nothing to do with dollars. Perhaps the added complexity has something to do with the nature of charities in a market society. Indeed, the reason charities are eligible to receive tax deductible contributions is that they are required to use these monies for charitable purposes, as opposed to mutual benefit purposes as is the case with other tax-exempt nonprofits. Thus, when the IRS grants an organization tax-exempt charitable status, it is not only excusing the organization from the requirement to pay federal income tax; it is also signaling to potential donors that the organization has given every indication that the donation will be used for charitable purposes and that the public can trust in this assurance.

\textbf{B. Benefits of the Exemption}

In addition to being excused from the requirement to pay tax on annual income, the charitable tax exemption opens doors to a variety of other economic and non-economic benefits. Among the economic benefits to a nonprofit corporation with tax-exempt charitable status are relief from the requirement to pay federal

\textsuperscript{81}I.R.C. § 6033 (West 2002) requires all exempt organizations to file an annual informational return (either IRS form 990 and 990T).

\textsuperscript{82}See. e.g., IRS form 1024 required for non-charities.

\textsuperscript{83}One example of the added complexity of Form 1023 as compared to Form 1024 is that Form 1023 asks more specific and detailed questions about the organization’s activities. On the other hand, Form 1024 allows the organization to write a narrative of its activities with only a few minimum requirements.
unemployment taxes,\textsuperscript{84} access to tax-exempt government bonds,\textsuperscript{85} and eligibility for preferred postal rates.\textsuperscript{86} In addition to these direct economic benefits, tax-exempt charities are also eligible for many indirect economic benefits granted by state and local governments. Some of these state and local economic benefits include exemption from state and local income taxes,\textsuperscript{87} state and local sales taxes,\textsuperscript{88} and local real property taxes.\textsuperscript{89} Thus, to the extent that a nonprofit corporation has economic motivations for seeking tax-exempt charitable status, those motivations may relate to the federal income tax exemption or, possibly, to one or more of the many other economic benefits.

Importantly, there are many non-economic benefits that accompany tax-exempt charitable status that might be more “valuable” than the various economic benefits. For example, the “halo effect” that accompanies charitable status may have no economic value at all. The “halo effect” is that indeterminable aspect of charitable status that results in positive emotional effects. These positive effects may exhibit themselves in many ways. Examples include when one gives money to a church without concern for receiving a charitable tax deduction, when one serves on a charitable board without expectation of financial payment for services or when the public just generally has a positive view of an institution due solely to the fact that it is mission-driven as opposed to profit-driven.

\textsuperscript{84}IRC § 3306(c)(8) (West 2002). Under the present system, charities only have to pay the equivalent of the former employee's unemployment compensation. Bazil Facchina, Evan A. Showell & Jan E. Stone, Privileges & Exemptions Enjoyed by Nonprofit Organizations, 28 U.S.F. L. Rev. 85, 102 (1993)[hereinafter Facchina].

\textsuperscript{85}IRC § 145(a)(1) (West 2002) (allowing state and local governments to issue bonds paying interest, exempt from federal income tax, to organizations described in § 501(c)(3)). Ninety five percent of the net proceeds from these tax-exempt bonds must be used by the charity and all property purchased with the bond proceeds has to be owned exclusively by the charity. Id.

\textsuperscript{86}The current postal regulations give religious, educational, scientific, philanthropic, agricultural, labor, veterans', and fraternal organizations second and third class nonprofit rates. Facchina, supra note 53, at 112. The only requirement is that the nonprofit mailers must be organized and operated for the primary purpose of the organization. Id. at 113.

\textsuperscript{87}For example, Pennsylvania provides an exemption to charities from state and local income taxes. 72 P.S. § 5020 – 204 (West 2005).

\textsuperscript{88}Illinois automatically exempts 501(c)(3) organizations for its sale/use tax. 35 ILCS 105/3-5 (West 2005).

\textsuperscript{89}California provides a property tax exemption to nonprofits. Cal Rev & Tax Code § 214.15 (West 2005).
Even though the halo effect produces real “value” - both for the corporation and for the public at large, that value is not often calculable in economic terms.

In addition to the halo effect, the charitable tax exemption allows for diversity and experimentation that often lead to production of undiscovered values. While these values may eventually be calculable, they are at first indeterminable and unknown. Finally, some aspects of the charitable activity are not reducible to dollars and cents and are, therefore, not capable of the types of comparison required for economic analysis. For example, economic values do not fully reflect the value of an education in an multi-racial environment as compared to a racially segregated environment. These ethical, emotional, moral and indeterminant values of charities are not something that can be fully explained in terms if economic efficiency, often because the value is shared by the charity and by those not associated with the charity. Thus, these non-economic values cannot be easily, or sometimes not at all, translated into economic terminology.

C. Obligations that Flow from the Exemption

Aside from the many benefits that accompany tax-exempt
charitable status, there are many obligations that a charitable corporation must abide by in order to obtain and maintain such status. Some of these obligations were previously alluded to in connection with the discussion of the requirement to file forms seeking charitable status and the requirement to annually report information concerning activities and operations. Statutorily, in order to obtain a charitable income tax exemption, a corporation must be:

organized and operated exclusively for religious, charitable, scientific, testing for public safety, literary, or educational purposes, or to foster national or international amateur sports competition . . . , or for the prevention of cruelty to children or animals, no part of the net earnings of which inures to the benefit of any private shareholder or individual, no substantial part of the activities of which is carrying on propaganda, or otherwise attempting, to influence legislation . . . , and which does not participate in, or intervene in . . . , any political campaign on behalf of (or in opposition to) any candidate for public office.

This statutory requirement has been described as imposing several affirmative and negative obligations on tax-exempt charities.

1. **Affirmative Obligations**

   a. “organized” and “operated”

   Affirmatively, tax-exempt charities must show that they are both “organized” and “operated” exclusively for a proper (or charitable) purpose. The “organized” requirement is met, quite straightforwardly, when the nonprofit applicant’s organizing documents and its federal forms used to apply for tax-exempt charitable status are in compliance with the law. Most notably, the organizing documents must clearly indicate that all assets of the proposed charity, upon dissolution, will go to other tax-exempt

---

93 See infra Part IIIA.
94 IRC § 501(c)(3) (West 2005).
95 E.g., by-laws and articles of incorporation.
96 E.g., IRS Form 1023
charities.\textsuperscript{98} Further, the organizational documents must state that all assets will be used in compliance with statutory requirements applicable to tax-exempt charities, namely for charitable purposes and not for inappropriate political, lobbying or private purposes.\textsuperscript{99} In other words, the organizing documents must affirmatively demonstrate that all charitable assets will remain in the charitable stream and that no profits will be inappropriately paid to private or political interests. The “operated” requirement is essentially the same as the organizational requirement, except that the operated requirement is concerned with whether the actual activities performed by the tax-exempt charity are consistent with the statutory requirements.\textsuperscript{100} Thus, not only must the nonprofit applicant say that it will be charitable, it must also demonstrate its charitableness by its actions.

\textit{b. proper (or charitable) purpose}

The proper purpose requirement is at the heart of the charitable tax exemption. This requirement imposes an obligation on the charity to have a special type of mission focus as opposed to a profit focus. The mission is not just any mission and is distinctly different from the mutual benefit mission of non-charitable nonprofit corporations.\textsuperscript{101} Instead, the mission that constitutes a proper purpose for the charitable tax exemption must be what is collectively referred to as a \textit{charitable} purpose. For simplicity, this Article uses the term “charitable” in a broad sense, referring to any variety of public purposes that might be acceptable under the statute.\textsuperscript{102} Some of these purposes are specifically delineated in the statute - religious and educational, for example. However, many charitable purposes have to be gleaned, by either the IRS or reviewing courts, from the statute. Some purposes that have been recognized as “charitable” include providing relief to the poor,\textsuperscript{103} protecting the environment,\textsuperscript{104}

\begin{itemize}
\item \textsuperscript{98}Treas. Reg. §1.501(c)(3)-1(b)(4) (2004).
\item \textsuperscript{99}Treas. Reg. § 1.501(c)(3)-1(b)(1) (2004).
\item \textsuperscript{100}Treas. Reg. § 1.501(c)(3)-1(c)(1) (2004).
\item \textsuperscript{101}A “mutual benefit” mission is one that is focused on primarily benefitting a particular group of persons who are oftentimes members. For example, a nonprofit social club or fraternity is a mutual benefit organization. \textit{See} Hansmann, \textit{supra} note 3, at 93-96.
\item \textsuperscript{102}I.R.C. § 501(c)(3) (West 2005).
\item \textsuperscript{103}Treas. Reg. § 1.501(c)(3)-1(d)(2) (2004)
\item \textsuperscript{104}Treas. Reg. § 1.501(c)(3)-1(d)(2) (2004); Rev. Rul. 76-204, 1976-1 C.B. 152.
\end{itemize}
February 24, 2006: DRAFT ONLY. DO NOT CITE/REPRODUCE/CIRCULATE WITHOUT PERMISSION

combating community deterioration, providing homes for the elderly, improving health and many more. All of these specifically recognized charitable purposes have at least one thing in common - they all seem to impart some “valued” benefit upon the public and not just a defined mutual benefit (or membership) group.

A key question for the charitable tax exemption is how to determine whether a particular purpose is “charitable” or not. True, a purpose that is arguably a subset of already recognized charitable purposes does not present many conceptual difficulties. For instance, if an organization has as its mission the protection of swamp land, that organization’s purpose is probably charitable because it involves protection of the environment in its natural state - a well-established charitable purpose. But there are at least three distinct ways in which determining whether a particular purpose is charitable or not becomes conceptually more complicated. The first instance is when a newly proposed purpose does not appear to fit within a subset of an already recognized charitable purpose. That is, what if, in order to conclude that a particular purpose is charitable, one would necessarily have to add another item to the list of potential charitable purposes. For example, would it have been “charitable” for Frederick Douglass to advocate that slavery was morally wrong during a time when slavery was completely legal and thought by many to be morally acceptable? Was it “charitable” for Nicolaus Copernicus to advocate that the sun, and not the earth, was the center of our universe at a time in history when all believed that the earth indeed was the center of the universe? The underlying conceptual quandary is not how these questions would be answered in hindsight, but rather, how should they be answered when they arise.

The second instance in which defining the term “charitable” becomes conceptually difficult is when the nature of a previously recognized charitable purpose has changed such that it no longer

---

109See Hansmann, supra note 3, at 57 ("[T]he repeated and unreflective reinterpretation of the exemption to accommodate new forms of nonprofit [charitable] activity, . . . . offers clear evidence of the lack of, and need for, a coherent policy on which to base the exemption.").
seems to be charitable. This conceptual difficulty is most prevalent today in the field of healthcare and the exemption of hospitals as “charitable” nonprofit corporations.\textsuperscript{111} Historically, hospitals were granted charitable tax-exempt status because they provided health care to the poor, an established charitable purpose.\textsuperscript{112} Over time, however, this view of the basis for exempting hospitals changed such that hospitals no longer are required to provide health care to the poor.\textsuperscript{113} Thus, we have a healthcare market in the United States in which for-profit hospitals operate along side charitable nonprofit hospitals.\textsuperscript{114} Does this state of affairs indicate that hospitals should no longer be tax-exempt because efficiency analysis suggests that hospitals would exist with or without the exemption?\textsuperscript{115} If so, is there some value aside from efficiency which would indicate that, despite the economics, we still need tax-exempt hospitals? Pursuant to contextual diversity, if the economic analysis suggest that hospitals should no longer be tax-exempt, an additional question to ask is whether there are values besides economic efficiency which would suggest continuing to exempt hospitals.

The third instance in which defining the term “charitable”

\textsuperscript{111}Hansmann, supra note 3, at 66-67. (arguing that many hospitals and other commercial nonprofits should not be tax-exempt as charities).

\textsuperscript{112}Rev. Rul. R 56-185, 1956-1 C.B. 202 (hospital must operate to the extent of “financial ability for those not able to pay for the services rendered” in order to sustain exemption under § 501(c)(3)).

\textsuperscript{113}Rev. Rul. 83-157, 1983-2 C.B. 94 (hospital does not have to operate emergency room open to the poor to sustain § 501(c)(3) tax exemption if “health planning agency has found that this would unnecessarily duplicate emergency services and facilities that are adequately provided by another medical institution in the community”).

\textsuperscript{114}The same dual tract phenomenon exists in other areas as well, such as homes for the aged. See e.g., Rev. Rul. 61-72, 1961-1 C.B. 188 (a home for the aged is exempt if it is dedicated to providing, among other things, care and housing to aged individuals who would otherwise be unable to provide for themselves without hardship); Rev. Rul. 72-124, 1972-1 C.B. 145 (concluding that, as an alternative to Revenue Ruling 61-72, “an organization . . . which devotes its resources to the operation of a home for the aged will qualify for charitable status . . . if it operates in a manner designed to satisfy the three primary needs of aged persons. . . . [:] the need for housing, the need for health care, and the need for financial security”).

\textsuperscript{115}See generally Hansmann, supra note 3. A related question might be: Does this state of affairs indicate that we should adopt a law which prohibits the operation of for-profit hospitals, leaving the entire market to the nonprofit sector? Although this question is just as important for economic analysis purposes, it is not as critical here because it does not directly bear on the issue of whether hospital operation is “charitable” and entitled to tax-exempt status.
becomes conceptually difficult is when a proposed purpose violates the public policy doctrine, a doctrine adopted by the Court and now incorporated into tax-exempt charity law. Pursuant to the public policy doctrine an organization that is otherwise “charitable” will not be eligible for charitable tax exemption if it engages in acts that contravene “clear” or “established” public policy. The prototypical example of an instance in which the public policy doctrine would defeat charitable status is racial discrimination against blacks. For example, the Supreme Court affirmed the IRS’ revocation of charitable status for a nonprofit religious school that discriminated against blacks in its admission policies.

The public policy doctrine, as applied to racial discrimination against blacks, is not problematic in the sense that it is nearly universally accepted that discrimination against a person because she is black is morally repugnant. However, the public policy doctrine becomes difficult to apply in instances other than invidious racial discrimination against blacks. For example, should the public policy doctrine be applied in such a way as to deny charitable status to nonprofit schools that make racial preferences in the context of affirmative action? While efficiency analysis might suggest that all racial preferences are prohibited by the public policy doctrine, how does this analysis accommodate the teachings of Critical Race Theory, which teaches that race based affirmative action is meaningfully different from invidious racial discrimination? Thus, the conceptual difficulty with the public policy doctrine is how does one determine the existence of “clear” or “established” public policies other than the policy against invidious racial discrimination?

Addressing these three instances of conceptual difficulty concerning the meaning of the term “charitable” - new purposes, old purposes and anti-public policy purposes - must lay at the heart of any theory of charitable tax exemption. If the charitable tax exemption has any normative coherence, it would seem that the coherence must

\[116\text{See Bob Jones Univ. v. United States, 461 U.S. 574, 591 (1983) (“A corollary to the public benefit principle is the requirement, long recognized in the law of trusts, that the purpose of a charitable trust may not be illegal or violate established public policy.”).}\]

\[117\text{Id. at 612.}\]

\[118\text{David A. Brennen, Race-Conscious Affirmative Action by Tax-Exempt 501c3 Corporations After Grutter and Gratz, 77 St. Johns Univ. L. Rev. 711, 725-730 (2003).}\]

\[119\text{See, infra Part IB, for discussion of critical race theory.}\]
somehow address these purposes since they collectively represent the areas in which the parameters of the exemption are most often tested. Such is the implicit assumption of all existing theories of charitable tax exemption. While many of the existing theories of charitable tax exemption rationalize portions of these conceptually difficult aspects of defining charitable, many do not. In many cases, the lack of rationalization might be due to the absence of concern for that which cannot be rationalized by efficiency analysis. But that is the point - efficiency analysis cannot rationalize or explain all that exists in law. And while some theories of the charitable tax exemption have gone beyond efficiency and have resorted to non-economic analysis to explain the charitable tax exemption, even these theories - arguably - do not go far enough.

Like the existing theories of charitable tax exemption, the contextual diversity theory espoused in this Article presumes that understanding the scope of the term “charitable” is key to understanding the normative rationale for the exemption. It is not enough to say that “charitable” simply means that which benefits the public. Where does that get us? How should we determine whether any particular activity benefits the public? Importantly, how should we determine whether a particular activity, that could benefit the public if done in a particular way, fails to do so because it contravenes “established public policy?” LMT, as explained earlier in this Article, suggests that we may not be able to define the outer limits of charity (public benefit) due to the dynamic, ever-changing and transformative nature of the marketplace. Thus, the best we can hope for is to continually re-evaluate, from a variety of perspectives, what actually provides a benefit to the public. While

---

120 See discussion of various theories infra Part IV.
121 Bittker & Rhadert, supra note 2, at 330-333.
122 See, e.g., Malloy, Framing the Market, supra note 11, at 63-64. Professor Malloy explains:

Law and market economy theory involves the study of the social/market exchange process by focusing on the relationship between law, culture, and markets. This relationship is triadic, dynamic and multi-directional. Moreover, in this relationship, one can understand the market sphere as expressing a concern for individualization, with a market focus on the pursuit of self-interest. On the other hand, culture is a collective concept, and therefore the cultural sphere can be understood as expressing a community perspective or a notion of the public interest.

Id.
many of these perspectives might involve economic analysis, many will not. Instead, “public benefit” may at times be determined based upon humanistic or other non-economic conceptions of justice, fairness, equality or other important values.

2. Negative Obligations

In addition to the many affirmative obligations that flow from tax-exempt charitable status, many negative obligations also exist. Among the negative obligations are limits on political campaigning, legislative lobbying, private inurement and private benefit.

a. politics and lobbying

The political campaign prohibition provides that tax-exempt charities cannot “participate in, or intervene in . . . , any political campaign on behalf of (or in opposition to) any candidate for public office.” Textually, this is an absolute prohibition in the sense that any amount of prohibited political campaign activity by a tax-exempt charity will result in the charity’s loss of tax-exempt status. However, the IRS has indicated that even political-looking activities that may otherwise be covered by the political campaign prohibition will not cause loss of exemption if the activities do not show a bias for (or against) a candidate. Additionally, the political campaign prohibition does not foreclose many activities often associated with political campaigns, such as voter education and “get out the vote” campaigns. Thus, the primary aim of the political campaign prohibition for charities is to prevent the usurpation of a charity for private, as opposed to public, political purposes.

The legislative lobbying limitation provides that “no substantial part of the activities” of a charity can consist of “carrying on propaganda, or otherwise attempting, to influence legislation (except as otherwise provided in subsection (h)).” This lobbying limitation effectively means that a charity cannot lobby elected officials so as to affect legislation in a particular (biased) way. Unlike the political campaign prohibition, the lobbying limitation is not a complete ban on legislative lobbying. Instead, the lobbying limitation

124 See Jones, The Tax Law of Charities, supra note 76, at 472.
only prohibits charities from engaging in more than an insubstantial amount of lobbying. Additionally, the lobbying limitation does not prohibit certain types of activities that clearly benefit the public, such as speaking at a legislative hearing concerning legislation pursuant to an invitation of the legislative officials.¹²⁶ Thus, much like the political campaign prohibition, the apparent aim of the lobbying limitation is to limit lobbying where it presumably fails to provide a public (as opposed to private) benefit.

Together, the political campaign prohibition and the legislative lobbying limitation serve to limit a charity’s involvement in government matters to the extent that little public benefit is likely to result. Arguably, these limitations on governmental influence might be required even without the textual expression of them in the statute, at least to the extent that the campaigning or lobbying either violates one of the other negative requirements for exemption¹²⁷ or causes a violation of the “exclusivity” requirement.¹²⁸ However, having the textual expression of these limitations in the statute gives force to their anti-public aspects. While the campaign prohibition and the lobbying limitation may indeed have economic rationales, they could just as well have non-economic political rationales. For example, it could be that these government involvement limitations are a reflection of the sovereignty view of the charitable tax exemption - that is, the view that charities are a type of sovereign akin to state and local government entities.¹²⁹ Accordingly, any theory which attempts a normative explanation for the charitable tax exemption must also address this government involvement limitation.

b. private inurement/benefit

Like the other negative obligations, the private inurement prohibition and the private benefit limitation have statutory origins. The private inurement prohibition provides that “no part of the net earnings” of a tax-exempt charity can “inure[] to the benefit of any private shareholder or individual.”¹³⁰ This prohibition has been a part of tax-exempt charity law in the United States since the first corporate

¹²⁷ See discussion of private benefit or private inurement prohibition infra Part IIIC2b.
¹²⁸ See Jones, The Tax Law of Charities, supra note 76, at 446.
¹²⁹ Brody, supra note 3, at 589.
tax exemption in 1909 and, in fact, was the only negative obligation for charities at that time. That original private inurement prohibition expressly prohibited distribution of the charity’s financial surplus to the charity’s controllers. Thus, the focus, at that time, was on preventing managers and others who controlled a charity from improperly taking profits from the charity.

The law regarding private inurement is structured in much the same way today in that the private inurement prohibition prevents certain “insiders” from taking charity profits other than as fair compensation for benefits conferred upon the charitable corporation. The prohibition is absolute in the sense that even if a “scintilla” of charitable profits go to insiders, the charitable exemption is lost. The term “insiders” is not defined in the statute except to state that charitable profits cannot go to any “private shareholder or individual.” The IRS and the courts have developed divergent interpretations of precisely who might be considered an insider. Under the IRS’ expansive view of the term, “insiders” are persons having a personal or private interest in the activities of the charity. Thus, insiders could include those who have actual control of the corporation (such as board members and managers) and those who have virtual or constructive control of the charity (such as employees, and even independent contractors). The courts, on the other hand, take a less expansive view of the phrase “insiders,” often limiting the phrase to board members and creators of the charity.

The private inurement prohibition clearly illustrates an important aspect of this Article’s thesis that economic analysis alone cannot fully explain the charitable tax exemption. Economic analysis only tells us that diversion of corporate profits from a charity to private interests (insiders) is prohibited. What economic analysis

\[131\text{See Jones, The Tax Law of Charities, supra note 76, at 306. See also Corporation Excise Tax of 1909, Ch. 6, § 38, 36 Stat. 11, 115 (1909). In the first corporate tax exemption law, the only expressly stated requirement was that any financial surplus derived from conducting the charitable endeavor could not be distributed to those who control the entity by which charity is delivered. See id.}

\[132\text{The legislative history for the first corporate tax exemption law provides that certain entities provide no person with a “scintilla of individual profit” from the entity. 44 Cong. Rec. 4150-51 (1909) (statement of Sen. Bacon). See also Darryll K. Jones, The Scintilla of Individual Profit: In Search of Private Inurement and Excess Benefit, 19 Va. Tax Rev. 575, 591 (2000).}

\[133\text{Treas. Reg. § 1.501(a)-1(c) (2004).}

\[134\text{United Cancer Council, Inc. v. Comm’r, 165 F.3d 1173, 1176-1177 (7th Cir. 1999).}
does not necessarily tell us how to determine who an “insider” might be. Should one take an expansive view of the term “insider,” which would severely limit certain aspects of permissible charity operations? Or should one take a narrow view, which may open up possibilities for charity operations that might not otherwise exist? Whichever view one takes, economic analysis alone does not necessarily drive the entire decision. For instance, if one has a perspective of law which views government involvement in private affairs as undesirable, such as the view taken by many conservatives, one might choose a narrow view of the term “insider.” On the other hand, if one has a perspective of law which welcomes government supervision of charities - a liberal view - one might choose the more expansive view of the term “insider.” Thus, one’s political perspective of law - either conservative or liberal, for example - might determine which view of the term “insider” is more valuable without regard to efficiency.

Although the private benefit limitation is also statutory based, it is not as explicitly spelled out in the statute as is the private inurement prohibition. Instead, the private benefit prohibition resulted from interpretation of the term “exclusively” as used in the charitable exemption statute. The IRS and various courts have determined that the term “exclusively” really means “primarily.” Thus, a tax-exempt charity will be considered exclusively charitable so long as it is primarily charitable. This means that a tax-exempt charity can perform a small amount of private benefit in addition to its primarily public benefit functions. In addition to not being an

---

135“Corporations, and any community chest, fund, or foundation, organized and operated exclusively for religious, charitable, scientific, testing for public safety, literary, or educational purposes.” I.R.C. § 501(c)(3) (West 2005) (emphasis added).


137One area in which this aspect of the private benefit limitation has caused controversy is in when a charity joins with a for-profit entity to form a partnership. See eg, Redlands Surgical Services v. Comm’r 113 T.C. 47 (1999); St. David’s Health Care Sys v. United States 349 F.3d 232 (5th Cir. 2003). The government and taxpayers disagree as to how the private benefit limitation should be interpreted in this context. For instance, the predominant government view is that the creation of the partnership causes loss of charitable status if the charity does not have control of the partnership. Rev. Rul. 98-15, 1998-1 C.B. 718; St. David’s Healthcare Sys. Charities, on the other hand, argue that the issue is not control, but rather whether the partnership actually functions as an exclusively charitable entity. In at least one jurisdiction (the 5th circuit), the court has determined that control is the appropriate
absolute prohibition, another difference between the private inurement prohibition and the private benefit limitation is that the private benefit limitation is not focused solely on “insiders.” According to the IRS, if a charity provides more than an insubstantial amount of private benefit to anyone - insider or not - it will no longer be considered “exclusively” charitable. Thus, despite the differences in terms of substantiality and prohibited beneficiaries, the private benefit limitation works in conjunction with the private inurement prohibition to ensure that charitable assets remain in the charitable (public) stream.

As with the other aspects of the charitable tax exemption, any normative theory should attempt to account for the private benefit and private inurement aspects of the exemption. Indeed, every theory espoused thus far has not only accounted for the private benefit and private inurement obligations, these obligations play a central role in every economic rationale for the charitable tax exemption.\textsuperscript{138} This is as it should be. For without the requirement that private benefit and inurement be minimized, it is difficult to comprehend any coherent explanation for the charitable tax exemption. Thus, to this extent, economic analysis does have true value as an analytic tool.

LMT teaches us that some aspects of law can and should be explained by economics.\textsuperscript{139} However, this does not mean that all, or even most, aspects of law can or should be so explained. Many aspects of law, including many aspects of tax-exempt charity law, have rationales that are not amenable to traditional neoclassical economic understanding or analysis. One cannot necessarily place an

---

\textsuperscript{138}cite examples from various theories.

\textsuperscript{139}Malloy, \textit{Framing the Market}, supra note 11, at 56-64.
economic value on notions of justice, fairness and equality that are not based on economic efficiency. Therefore, the key to understanding the charitable tax exemption is recognizing that any theoretical rationalization of the exemption must be broad based and not rely only on notions of positive economic efficiency. In order to better understand this idea, Part I of this Article explains how economic theories can be used effectively with non-economic theories to both explain and sculpt law. LMT articulates this idea and Professor Emily Houh has demonstrated its application in understanding the good faith exception in contract law. Now, a few words about commercial competition of charities with for-profit corporations and private foundation charities.

D. Commercial Competition

Even though they are mission-focused and not profit-focused, charities often engage in commercial profit making activities with the aim of furthering a distinct charitable purpose. One common by-product of this manner of operation is that charities sometimes compete with for-profit firms for market dollars. This competition - actual or constructive - might result from either of two possible circumstances. The first is when the charity engages in a commercial activity for the sole purpose of raising money to advance its charitable mission. A second circumstance of competition is when the charity engages in a commercial activity in such a way that the activity itself (as opposed to the revenues generated by the activity) furthers or constitutes its charitable purpose. In an effort to make the charitable tax exemption more effectively focused on income generated by charitable activity, two developments in the law occurred. The first development was repeal of the destination of income test for charitable exemption - thus making charitable activity itself more important than how profits are used. The result was to prohibit charities from engaging in more than an insignificant amount of non-charitable commercial activity with the aim of raising money for a charitable purpose. The second development occurred in Congress with the adoption of the unrelated business income tax in

140 See Houh, supra note 22.
142 Id.
143 See I.R.C. § 502(a) (denying § 501(a) exemption to feeder organizations).
1950, with later revisions in 1969. The institution of the unrelated business income tax allows the law to treat nonprofit charities like all other types of corporations to the extent that its activities are not mission-focused and produce income. Thus, with these two legal developments, the federal income tax exemption no longer applies to profits generated by a nonprofit charity if the activities generating these profits are not themselves charitable.

As with many aspects of the charitable tax exemption, the basic idea of dealing with unfair commercial competition by eliminating the destination of income test and adopting the unrelated business income tax is surely efficiency rationalized. Indeed, the addition of these legal rules to tax-exempt charity law advance the notion that only the income from activity that is charitable will be exempted from the federal income tax. However, beyond this basic idea, there are other issues to contend with concerning commercial competition that cannot be resolved by economic efficiency analysis alone. For instance, in the late 1990's the IRS had to decide whether and when to treat payments made to charities by charitable event sponsors as tax-exempt revenues or as taxable unrelated business income tax revenues. Many aspects of these corporate sponsorship issues were resolved not by economic analysis, but by other non-economic perspectives of law. Thus, as is the case with many aspects of law, economics can only go so far when it comes to a theoretical rationalization for the charitable tax exemption. This means that even economically motivated laws, such as tax law, could and should be understood not only in terms of economics, but also in terms of non-economic conceptions of justice and fairness. Part I of this Article demonstrates more explicitly how economic analysis can be used with other theories of law to explain and shape law.

E. Private Foundations

---

146 A tax is imposed on the unrelated business income of a tax-exempt organization. I.R.C. 511 (West 2002). Unrelated business taxable income is any gross income derived by an organization from any unrelated trade or business. I.R.C. 512 (West 2005). Unrelated trade or business is any trade or business that is not substantially related to the exercise or performance of that organization’s exempt purpose. I.R.C. 513 (West 2002).
148 Need example here.
Although many charities receive broad-based financial support from government and non-governmental entities, many other charities do not have such varied and wide-spread sources of financial support. When the financing sources for a charity are concentrated in one or very few people - people who might also have direct or indirect control over the charity - the potential for abuse of tax-exempt status is more likely to occur than if the sources of financial support and control are more widely dispersed. In recognition of this reality, Congress enacted rules in 1950 that denied tax exemption to charities that engaged in “prohibited transactions,” unreasonably accumulated income, were used substantially for non-exempt purposes, or had invested money so as to jeopardize exempt purposes.\(^4\) Congress exempted churches, schools and colleges, hospitals and certain publicly supported organizations from these new rules.\(^5\) These 1950 rules sparked the beginning of a distinction in tax-exempt charity law between “public” charities and “private” charities (private foundations) that continues today in a modified form.

In 1954 and 1964, Congress made additional legislative changes to further distinguish so called “public” and “private” charities by providing for an increased ceiling on charitable contribution deductions for contributions to “public” charities.\(^6\) These “public” charities included churches, schools, hospitals and “publicly and governmentally supported” charities. Finally, in 1969, Congress again acted to stem the tide of potential abuses by “private” and so-called non-operating charitable foundations.\(^7\) In that year, Congress eliminated the rules that denied charitable tax exemption to

---

\(^4\)See e.g., Revenue Act of 1950, 81 P.L. 814, 64 Stat. 906, 957-59 (1950). The Revenue Act of 1950 denied income tax exemption to private foundations that “engaged in a prohibited transaction.” Id. The term “prohibited transaction” was defined as including a transaction in which a charity “lends any part of its income or corpus,” “pays any compensation, in excess of a reasonable allowance,” “makes any part of its services available on a preferential basis,” “makes any substantial purchase of securities . . . for more than adequate consideration,” “sells any substantial part of its . . . property, for less than an adequate consideration in,” or “engages in any other transaction which results in a substantial diversion of its income or corpus” to “the creator”; substantial contributor; a family member of a creator or family member; or a corporation controlled by a creator. Id.


\(^7\)See I.R.C. § 4941-4945 (West 2002).
these troublesome charities and replaced those rules with new excise taxes that act as penalties for charities that engage in potentially abusive behavior, including self-dealing, unreasonable income accumulations, excess business holding, risky investments, and disfavored expenditures.

The essential role of the private foundation rules is to control nonprofit corporations that express a desire to be mission-focused, but that lack a public mandate that would support the mission. For example, assume a wealthy individual has a family member who contracts a rare incurable disease. If neither the government nor any existing private organization is currently searching for a cure for that disease, the wealthy individual could either fund the research on her own, or she could donate the money to a self-created charity whose sole mission is to find a cure. While this self-created charity is most certainly mission-focused, the fact that neither the government nor any other private group is researching this disease indicates a lack of a public mandate for finding a cure. The lack of a public mandate, however, does not mean that the effort at finding a cure is not a proper charitable purpose. Nonetheless, the lack of widespread public involvement in this effort may lead to financial abuses of the self-created charitable corporation in terms of investments and the like. The private foundation rules are intended to hone in on the areas of potential abuse and, thereby, protect the charitable fisc. The heart of the private foundation rules is the definition of the term “private foundation.”

A private foundation is a charitable corporation that either has a wide array of public support or public patrons. Thus, the proxy for the public mandate is the presence of a wide array of financial support and supporters. In some cases, the charitable purposes themselves (e.g., churches, hospitals and schools) serve as the proxy.

The presence of the private foundation rules is consistent with contextual diversity because the rules permit all sorts of purposes to

---

154 I.R.C. § 4942 (West 2002).
155 I.R.C. § 4943 (West 2002).
156 I.R.C. § 4944 (West 2002).
158 I.R.C. § 509(a) (West 2002).
be advanced by tax-exempt charitable status - not only purposes that many people agree are of benefit to the public or are willing to fund, but also purposes that very few believe might benefit the public or are willing to fund. To demonstrate the benefit of private foundations, consider the example of the Howard Hughes Medical Institute - a private foundation that, at nearly $11 billion, is the second richest private foundation in the country.\textsuperscript{162} The foundation originated from the personal fortune of one man, Howard Hughes, who created the foundation in the early 1950's as a personal tax shelter to protect his fortune. The stated purpose of the foundation is to do medical research. Although the foundation did no research for a number of years, its research efforts have since resulted in discovery of the genes responsible for cystic fibrosis and muscular dystrophy, a non-invasive test for colon cancer and a drug that fights leukemia. More recently, the Howard Hughes Medical Institute has created new stem cell lines for future medical research - something that the federal government is prohibited from doing due to a Presidential directive.\textsuperscript{163} Without this privately funded charity, these many valuable discoveries might not have occurred.

While diversity means that we have these opportunities for creativity that might not otherwise exist, they must be considered in context. Given the private nature of the funding source and the incentive for tax law abuse, the private foundation rules serve the governmental/public purpose of stemming the likelihood of tax abuse while still permitting the substantive charitable functions to take place. Thus, the private foundation rules are a recognition of the fact that even private individuals can create great wealth and value for society. These rules are also a recognition of the fact that government, as a proxy for the public, has a strong interest in not allowing individuals to abuse the charitable tax exemption. In essence, the private foundation rules reflect the diverse ways in which public benefit might be “discovered.” Diversity is a very important aspect of the charitable tax exemption. Contextual diversity is important because it recognizes the important role of diversity while also allowing charity tax law to mediate or modulate diversity in

\textsuperscript{162}60 Minutes: Howard Hughes Medical Institute (CBS television news broadcast, Nov. 23, 2003) (transcript on file with Burrelle’s Information Services).

order to arrive at the right mix of public interests and private interests represented in the charitable tax exemption. This mediating aspect of the charitable tax exemption must accommodate a variety of value interests - both economic and non-economic value interests. Part I of this Article explains in greater detail how economic values can be effectively combined with other non-economic values to form a more complete theoretical understanding of law and, potentially, aid in better refining law in the future.

IV. **EFFICIENCY THEORIES OF CHARITABLE EXEMPTION**

Traditional theories of the charitable tax exemption - at least those promulgated prior to 1990 - are principally based on concepts of economic efficiency. These efficiency based theories explain the charitable tax exemption as either a subsidy by government for public goods, a necessary result of using net income to define tax liability or a means of compensating charities for capital constraints. Other efficiency based theories contend that the charitable tax exemption is either a payment for an entity’s ability to garner donative support or a means of compensating charities for the risk they assume in providing public goods. Each of these economic theories for the charitable tax exemption has its strengths and its weaknesses. In terms of explaining the economics for why charities are tax-exempt, these traditional theories do a pretty good job. These theories are also somewhat useful in sculpting the contours of the charitable tax exemption. However, these traditional theories lack significant non-economic components which, ultimately, makes them incomplete. These economic theories for the charitable tax exemption do not explain the existence of the many non-economic aspects of the exemption. This explanatory deficiency also means that these efficiency theories cannot fully guide us in sculpting the contours of charitable tax exemption law.

A. **Public Benefit Subsidy Theory**

---

165 Bittker and Rahdert, *supra* note 2.
166 Hansmann, *supra* note 3.
The first of the efficiency based theories of charitable tax exemption is the public benefit subsidy theory. The public benefit subsidy theory posits that the charitable tax exemption is a means by which government encourages organizations engaged in providing public goods to continue to do so. The most notable proponent of this theory has been the government itself. For example, in Bob Jones University v. United States, the Supreme Court notes the following:

Charitable exemptions are justified on the basis that the exempt entity confers a public benefit -- a benefit which the society or the community may not itself choose or be able to provide, or which supplements and advances the work of public institutions already supported by tax revenues. History buttresses logic to make clear that, to warrant exemption under § 501(c)(3), an institution must fall within a category specified in that section and must demonstrably serve and be in harmony with the public interest. The institution's purpose must not be so at odds with the common community conscience as to undermine any public benefit that might otherwise be conferred.\(^{169}\)

In addition to the government, scholars have also advocated this public benefit subsidy theory for the charitable tax exemption. An essential aspect of the theory is that government subsidizes certain "goods" or services that government either cannot or will not supply on its own. The reasons for government failure to supply these particular goods or services vary. For instance, government could have constitutional constraints - as is the case with religion - that prevent government from supplying the good. Government might also have political constraints, such as the requirement for majority political support, that prevent it from supplying the good. Thus, the point of the charitable tax exemption - according to the public benefit subsidy theory - is to permit government to essentially “pay” or “compensate” private entities that supply these public goods and services.

An implicit assumption underlying the public benefit subsidy theory is the idea that government, under neutral principles, can

\(^{169}\)See Bob Jones Univ. v. United States, 461 U.S. 574 at 591-92.
determine what constitutes a public good or service.\footnote{See Atkinson, \textit{Altruism supra} note 3, at 606.} Hence, elements of the rational market participant - reminiscent of economic analysis - pervade this theory of charitable tax exemption. Also of economic dimension is the idea implicit in the public benefit subsidy theory that the government is somehow \textit{paying} charities for what they produce.\footnote{See \textit{Bob Jones Univ. v. United States}, 461 U.S. 574 at 590: The exemption from taxation of money or property devoted to charitable and other purposes is based upon the theory that the Government is compensated for the loss of revenue by its relief from financial burdens which would otherwise have to be met by appropriations from other public funds, and by the benefits resulting from the promotion of the general welfare. (citing H. R. Rep. No. 1860, 75th Cong., 3d Sess., 19 (1938)).} Given these economic dimensions, the traditional public benefit subsidy theory partially rationalizes many aspects of the charitable tax exemption. That is, the exemption truly is a form of financial support for charities - at least in some cases. Further, the government does play a role in deciding what goods and services actually benefit the public for purposes of the charitable tax exemption. Nevertheless, there is much about the charitable tax exemption that the traditional subsidy theory does not address.

For all of its great virtues, the public benefit subsidy theory does not address why this government financial support for charities must take the form of a tax exemption. For instance, why not simply have the government make direct grants to nonprofit corporations that provide goods and services that benefit the public? The public benefit subsidy theory also does not articulate a coherent rationale for how the decision is made as to what goods and services benefit the public. The theory’s silence on this point seems to indicate that neutral market principles might drive the process of deciding what benefits the public. However, this is not necessarily so. In \textit{Bob Jones University}, the private university that was the subject of that case provided what was identified in the charitable tax exemption statute as a public benefit - education; yet the Court held that the education in that case was not entitled to exemption due to the presence of invidious racial discrimination.\footnote{Justice Powell, though he agreed with the majority’s ultimate holding that charities could not maintain tax exemption if they engage in invidious racial discrimination, questioned the broad majoritarian nature of the majority’s adoption of the public policy doctrine to address this specific racial subordination problem:}

\footnote{\textit{See Atkinson, \textit{Altruism supra} note 3, at 606.}}
\footnote{\textit{See Bob Jones Univ. v. United States}, 461 U.S. 574 at 590: The exemption from taxation of money or property devoted to charitable and other purposes is based upon the theory that the Government is compensated for the loss of revenue by its relief from financial burdens which would otherwise have to be met by appropriations from other public funds, and by the benefits resulting from the promotion of the general welfare. (citing H. R. Rep. No. 1860, 75th Cong., 3d Sess., 19 (1938)).}
\footnote{Justice Powell, though he agreed with the majority’s ultimate holding that charities could not maintain tax exemption if they engage in invidious racial discrimination, questioned the broad majoritarian nature of the majority’s adoption of the public policy doctrine to address this specific racial subordination problem:}
does not provide a rationalization for this aspect of charitable tax exemption. Hence, the public benefit subsidy theory’s reliance on “neutral” efficiency principles simply does not provide a basis for understanding how a public benefit is determined.

B. Base-Defining Theory

Recognizing deficiencies in the public benefit subsidy theory, Boris Bittker and George Rahdert developed a theory of charitable tax exemption that also focused explicitly on the economic aspects of the exemption. Bittker and Rahdert’s base-defining theory posits, essentially, that charities (and many other nonprofits) are exempt from the federal income tax because they are not suitable targets of the income tax. Specifically, Bittker and Rahdert state that:

... [Charities] should be wholly exempted from income taxation, because [(1)] they do not realize “income” in the ordinary sense of that term and because, [(2)] even if they did, there is no satisfactory way to fit the tax rate to the ability of the beneficiaries to pay.

Thus, according to Bittker and Rahdert, charities are exempt from the income tax by necessity.

In support of their theory, Bittker and Rahdert explain that measuring the income of a charity is a conceptually difficult, if not impossible, task. To begin with, measuring an entity’s income requires a determination of the entity’s gross income in excess of expenses incurred in acquiring the income. Gross income is generally any economic enrichment that is not excluded from income by Congress. One common Congressional exclusion from income is gifts. Thus, money or property given with “detached and
disinterested generosity” is not usually treated as taxable income.\(^{179}\)

In looking at what a charity’s typical revenues might be (interest on endowment funds, membership dues, gifts/donations), Bittker and Rahdert conclude that, with the exception of interest on endowment funds, charities simply do not produce revenues that represent the types of enrichments that constitute taxable income.\(^{180}\) Bittker and Rahdert explain this conclusion from two perspectives. On the one hand, membership dues, gifts and donations to the charity would likely be viewed as excludable gifts from members/donors to the charitable entity. Even if not viewed as gifts to the institution, these charitable revenues might be viewed as excludable gifts to the charity’s beneficiaries - the charitable entity itself acting as a mere conduit for passing the revenues to beneficiaries.

In addressing the expense side of the net income equation, Bittker and Rahdert explain that, even if one were to properly conclude that charitable revenues constituted gross income, a separate difficulty involves determining what to count as deductible expenses incurred in acquiring the income.\(^{181}\) Bittker and Rahdert identified charitable expenditures as potentially including items such as staff salaries and medical welfare programs for indigents. One way of deducting an expense is by positioning it as an “ordinary and necessary expense incurred in carrying on a trade or business” activity.\(^{182}\) Bittker and Rahdert explain, however, that treating charitable activity as a “trade or business” is self-contradictory because, unlike for-profit enterprises, charities are mission focused, not profit focused. Additionally, even if the definition of “business” were expanded to include the business of providing charitable benefits, a charity would essentially end up having no tax liability because, as a result of the non-distribution constraint, all revenues must be devoted to mission purposes and no revenues may go as profits to insiders. Thus, net income - save for some instances of multi-year accumulations for specific purposes - would always equal zero, resulting in no tax liability. Another way of deducting expenses is by positioning the expense as eligible for the charitable contribution deduction.\(^{183}\) However, Bittker & Rahdert explain that, as with the business expense scenario described above, either

\(^{180}\)Bittker and Rahdert, supra note 2, at 307-309.
\(^{181}\)Bittker and Rahdert, supra note 2, at 309-314.
\(^{182}\)I.R.C. § 162(a) (West 2005).
\(^{183}\)See § 170(a).
structural impediments in the statute authorizing the charitable deduction or the necessary zeroing out of income that would result from allowing the deduction indicate that charities should not be permitted to take a charitable contribution deduction for charitable expenses.

In addition to the income measurement problems associated with imposing an income tax on charities, Bittker and Rahdert also raise concerns about the appropriate tax rate to apply to charities - further supporting their base-defining rationale for the charitable tax exemption.\textsuperscript{184} According to Bittker and Rahdert, tax rates are important because they implicate conceptions of efficiency related to either the “benefit” or “ability to pay” theories of taxation.\textsuperscript{185} The idea here is that tax law generally attempts to match tax burden with the taxpayer’s circumstances. Bittker and Rahdert argue that a charity’s income should be imputed to its beneficiaries for rate determination purposes since it is most likely the beneficiary who would bear the burden of any tax on the charity’s income. The problem, Bittker and Rahdert explain, is that the beneficiaries are usually unknown at the time the income is received and, thus, it is nearly impossible to determine an appropriate income tax rate. Even if the entity were taxed as a surrogate for the beneficiaries, Bittker and Rahdert explain that not knowing who the beneficiaries are would necessarily mean that a tax on income would be inefficient - potentially over taxing some beneficiaries and under taxing others. Bittker and Rahdert further explain that, aside from the identification-of-beneficiary aspect of the rate issue, another point is that these beneficiaries - were they to receive these charitable benefits directly - would be able to exclude them from income as excludable gifts.\textsuperscript{186} Thus, however the matter is approached, Bittker and Rahdert conclude that there is simply no way of coming up with a proper tax rate if charities were to be subject to the income tax.

As an economic explanation of the charitable tax exemption, Bittker and Rahdert’s base-defining theory does a good job of demonstrating that, for the most part, imposing an income tax on charities would likely not yield much in the way of federal income tax revenues. However, the thesis of this Article is premised on the notion that there is more to the charitable tax exemption than just the

\textsuperscript{184}Bittker and Rahdert, supra note 2, at 314-316.
\textsuperscript{185}Cite sources
\textsuperscript{186}See § 102(a) (West 2005).
elimination of a financial obligation of charities to pay tax on income - however that term is defined.\textsuperscript{187} Indeed, the charitable tax exemption is about justice and fairness in resource allocation; it is about providing and creating opportunities for societal enhancement and betterment where none would exist otherwise. Thus, Bittker and Rahdert’s theory, even if taken at face value, fails to fully explain the many non-economic aspects of the charitable tax exemption. More precisely, some of the aspects of the charitable tax exemption that Bittker and Rahdert’s theory fails to address include the difference between a zero or near-zero tax liability and tax exemption,\textsuperscript{188} political activities and lobbying,\textsuperscript{189} the definition of “charitable,” and private foundation rules.

Throughout their base-defining theory, Bittker and Rahdert explain that, even if the federal income tax were to apply to a charity’s income, it is quite likely that no tax revenue would result.\textsuperscript{190} Reminiscent of this view is the statement that “Tax immunity would then have been achieved, but by a more roundabout route than the straightforward exemption. . .”\textsuperscript{191} This view of the charitable tax exemption as nothing more than elimination of a financial obligation, completely overlooks the many other non-economic aspects of the charitable tax exemption. As Evelyn Brody explains quite well in her sovereignty theory of charitable tax exemption:

> While most observers have described tax exemption as a subsidy, a zero rate of tax differs qualitatively, not just quantitatively, from a one-percent rate of tax. Tax exemption maintains an independent distance between charities and the state. Similarly, exemption differs in an important political way from an equivalent system of direct grants.\textsuperscript{192}

Thus, in Brody’s words, there is a “qualitative[]” dimension to the charitable tax exemption that is not captured by a pure dollars and cents analysis. This qualitative difference is what lies at the heart of the normative justification for the exemption.

\textsuperscript{187}See Part III.
\textsuperscript{188}See Bittker and Rahdert, supra note 2, at 313.
\textsuperscript{189}Id. at 305, 334.
\textsuperscript{190}See Bittker and Rahdert, supra note 2, at 313.
\textsuperscript{191}Id at 313.
\textsuperscript{192}Brody, supra note 3, at 592-93.
As previously explained, central to the charitable tax exemption is defining the term “charitable.” Though Bittker and Rahdert address the issue of what constitutes “charitable,” they fail to fully develop the non-economic aspects of their theory as it relates to the meaning of “charitable.” For instance, when addressing the relevance of racial discrimination to the concept of “charitable,” Bittker and Rahdert explicitly minimize the importance of this relevance by referring to the race issue as a “minor problem[] of interpretation.” Granted, Bittker and Rahdert drafted their base-defining theory in 1976, well before the all important *Bob Jones University* case was decided in 1983. However, even in 1976 the predominant view was that racial discrimination rendered some purposes non-charitable because of racial discrimination’s inconsistency with federal public policy. The other “minor problem[] of interpretation” issue identified by Bittker and Rahdert also involved an issue of paramount importance to people of color: whether charities have an obligation to provide free or reduced cost services to those who are unable to pay in a variety of contexts. Could a critical race perspective add to our understanding of this aspect of the charitable tax exemption?

What’s important here is that Bittker and Rahdert’s base-defining theory fails to explain this and other non-economic aspects of the term “charitable.” Instead of recognizing this as a limitation of their base-defining theory, Bittker and Rahdert choose to minimize the non-economic issues as “minor.” Importantly, it is not only in the context of race, or even with regard to defining charitable, that Bittker and Rahdert must account for various aspects of the charitable tax exemption in some non-economic way. For example, with regard to private foundations, Bittker and Rahdert’s base-defining theory could not rationalize why the various private foundation excise taxes exist. Thus, the architects of the base-defining theory resort to a type of contextual diversity as a means of rationalizing these special

---

193 See Part III.
195 Id. at 331.
198 See Bittker and Rahdert, *supra* note 2, at 331 (n. 82). See also, Eastern Kentucky Welfare Rights Organization v. Simon (hospital); RR 74-587 (lender to “ghetto” businesses); RR 75-75 (public interest law firms).
penalty taxes. In rationalizing the private foundation excise tax rules, Bittker and Rahdert state that:

Private organizations displaying independence, flexibility, and originality are bound to tread on toes, and when the toes belong to public officials, an adverse legislative reaction should not come as a surprise. 200

Rationalizing the private foundation excise taxes as a type of penalty imposed for contravening government “territory” or “authority” seems consistent with this Article’s notion of contextual diversity. That is, law and market economy theory states that market participants are constantly seeking to gain authoritative control in the marketplace. 201

Bittker and Rahdert’s base-defining theory uses a similar type of non base-defining (non-economic) analysis to fully account for the educational exemption for museums, colleges and orchestras. 202 Because their beneficiaries are not necessarily poor, as is the case with many other types of charities, the improper rate aspect of the base-defining theory does not account for these particular types of “educational” institutions. 203 So, instead of relying exclusively on base-defining/economic concepts to explain these upper echelon charities, Bittker and Rahdert again resort to a type of contextual diversity analysis. Accordingly, Bittker and Rahdert rationalize that the benefits of “educational” institutions extend beyond the immediate beneficiaries to “an indefinably wide audience over the entire income spectrum.” 204 Additionally, they explain:

[I]t is precisely in the area of education, including the arts, that private institutions are especially well suited to serve as independent centers of power and influence in our society, fostering innovation and diversity with a dedication that government agencies can seldom muster or sustain. This separate rationale for tax exemption applies particularly to educational

200 Bittker and Rahdert, supra note 2, at 342.
201 See Malloy, LAW AND MARKET ECONOMY, supra note 13.
202 Bittker and Rahdert, supra note 2, at 334 and 335.
203 Cite language from B&R article.
204 Bittker and Rahdert, supra note 2, at 334.
Thus, when economics fails to explain some important and varied aspects of the charitable tax exemption, Bittker and Rahdert resort to notions of “diversity” and “context” to fill in the theoretical gaps.

### C. Capital Formation Subsidy Theory

Five years after publication of Bittker and Rahdert’s base-defining theory of charitable tax exemption, Professor Henry Hansmann published his own theory of charitable tax exemption which responds, in explicit economic terms, to Bittker and Rahdert’s approach. In his capital formation subsidy theory of tax exemption, Professor Hansmann explains that the rationale for the charitable tax exemption concerns the access of charities to capital markets. As Professor Hansmann explains:

... [T]he best justification for the exemption is that it helps to compensate for the constraints on capital formation that nonprofits commonly face, and that such compensation can serve a useful purpose, at least for those classes of nonprofits that operate in industries in which, for various reasons, nonprofit firms are likely to serve consumers better than would profit-seeking firms.

Thus, for Professor Hansmann, the tax exemption compensates charities for the lack of access to capital markets. Further, Professor Hansmann explains, this so-called “capital subsidy” promotes “efficiency when employed in those industries in which nonprofit

---

205 Bittker and Rahdert, supra note 2, at 335.
206 Actually, Professor Hansmann’s theory is a theory of the exemption of nonprofits generally - both charitable and non-charitable nonprofits. Hansmann, supra note 3, at 57 (referring to all tax exemptions listed in IRC 501(c), not just IRC 501(c)(3)). But since his theory includes the charitable tax exemption in its scope, this Article will refer to it as a theory of charitable tax exemption.
207 Id. at 56 (“Much of the discussion in this article is presented in economic terms, as is appropriate for the subject at hand.”).
208 In fact, Professor Hansmann refers to Bittker and Rahderts base defining theory as “the most comprehensive and thoughtful” of the efforts to rationalize the charitable tax exemption. Hansmann, supra note 3, at 54-55.
209 Id. at at 55.
firms serve consumers better than their for-profit counterparts.”\footnote{10}

Central to Professor Hansmann’s capital subsidy theory is the notion of contract failure. For Professor Hansmann, contract failure is a type of market failure that “derives from the inability of some or most consumers to make accurate judgements concerning the quality, quantity, or price of services provided by alternative producers.”\footnote{11} Contract failure, according to Professor Hansmann, is most prevalent with what he terms donative nonprofits (nonprofits that receive revenues mostly through donations) as opposed to commercial nonprofits (nonprofits that receive revenues mostly through commercial sales activities). Professor Hansmann’s classic example of a donative nonprofit that typifies contract failure is the American Red Cross.\footnote{12} A person making a contribution to Red Cross is in essence buying disaster relief services from the Red Cross for some unknown third party. This is a circumstance of contract failure because the consumer/donor must blindly rely on Red Cross to determine who gets disaster relief, how much they get, and under what terms they get it. Nonprofit firms are more efficient, for Professor Hansmann, than for-profit firms in providing these types of contract failure services because of the nondistribution constraint. That is, consumers are not as concerned with nonprofit firms as they would be with for-profit firms about donations being diverted to shareholders because nonprofit firms do not have shareholders. Thus, for Professor Hansmann, (donative) nonprofit firms are more efficient than for profit firms in circumstances of contract failure.\footnote{13}

In addition to contract failure, Professor Hansmann also points to constraints on the ability of nonprofits to obtain capital as another important component explaining the income tax exemption. The three major sources of funding for nonprofits are debt, donations and retained earnings.\footnote{14} Notably, because nonprofits cannot issue shares, they do not have access to equity capital, as is the case with for-profit firms. Professor Hansmann explains that debt capital is difficult for

\footnotetext[10]{Id. at 72.} \footnotetext[11]{Id. at 67-68.} \footnotetext[12]{Id. at 61.} \footnotetext[13]{According to Professor Hansmann, two efficiency advantages of donative nonprofits include “(1) a reduction in the efforts that consumers feel impelled to make to police the provider of a service when the provider is nonprofit rather than for-profit, [and] (2) a reduction in the disparity between cost and price occasioned by the elimination of the excessive profits that for-profit producers might be able to secure.” Id. at 70 (n.57).} \footnotetext[14]{Id. at 72-75.}
nonprofits to obtain because of the risk involved in loaning to nonprofits. Donations are also problematic because donations are uncertain and inadequate. Thus, according to Professor Hansmann, nonprofits must rely almost exclusively on retained earnings in order to finance growth. While this fact alone - restraints on access to capital markets - does not justify the tax exemption, Professor Hansmann argues that coupling this restraint with the fact that many nonprofits operate under circumstances of contract failure, means that the exemption is needed. For Professor Hansmann, if we want markets to operate at optimal efficiency, and if we accept that nonprofits are the most efficient producers of contract failure goods/services, then it makes sense that we subsidize nonprofits in order to increase the rate at which nonprofits can expand.215

Although Professor Hansmann’s capital formation subsidy theory articulates a clear rationale for why the charitable tax exemption is efficient, it does not articulate a clear basis for understanding aspects of the exemption that have no necessary connection to economic efficiency. Professor Hansmann’s deficiency is most apparent in his conclusion that the tax exemption should be sculpted so as to deny the exemption to many commercial nonprofits that produce simple standardized services, as opposed to complex services.216 Typical of Professor Hansmann’s view is the statement that “There would obviously be little point . . . in granting the exemption to a nonprofit hardware store.”217 What Professor Hansmann misses here is the point that even a hardware store might provide the type of benefit, under certain circumstances, that society wants, needs or otherwise values. For example, what if that hardware

\[ \text{id. at 74-75.} \]
\[ \text{id. at 86-89. Professor Hansmann explains:} \]
\[ \text{Between these two extremes-donative nonprofits on the one hand, and commercial nonprofits that provide simple standardized services on the other-we have the troublesome category of commercial nonprofits that provide complex personal services such as education, hospital care, nursing care, and day care. For which, if any, of these services are the fiduciary qualities of the nonprofit form so effective and necessary that tax exemption can be justified on efficiency grounds? It is difficult to offer an authoritative answer to this question, since at present there exist little solid data concerning the relative performance of nonprofit and for-profit firms in providing such services.} \]
\[ \text{id. at 88.} \]
\[ \text{id. at 87.} \]
store only employed people who are handicapped or blind? What if this hardware store provided an employment opportunity to racial minority groups or others who would not otherwise have employment? If for-profit firms choose not to open a hardware store that employs these populations, these people might be jobless or dependents of government. Thus, even though the hardware store might not operate under conditions of classic contract failure, and even though it might not be economically efficient to operate a hardware store by employing these populations, it is still of real value to society that this hardware store operate. To the extent that granting tax exemption allows this happen, then society is all the better for it.

Another example of Professor Hansmann’s theory’s disconnect from the aspects of the charitable tax exemption that are not susceptible to efficiency analysis is the assertion that nonprofit hospitals should not be eligible for tax exemption. Professor Hansmann’s articulated reason for this assertion is the lack of contract failure or need for capital evident in the hospital industry. It is important to realize - and this is a point that Professor Hansmann and many others miss - that the value inherent in a particular form of charitable organization may not be readily apparent by means of traditional efficiency analysis. To illustrate, consider Professor Jill Horowitz’s empirical research concerning hospitals. Professor Horowitz concludes that - despite the myriad of calls for ending tax exemption for hospitals that do not serve the poor - empirical research shows that tax exempt nonprofit hospitals provide societal benefits

\[\text{See, e.g., Hansmann, supra note 3, at 89. Professor Hansmann explains:} \]

\[\text{On the other hand, it is not at all clear that there is justification for the relatively recent decision to exempt nonprofit hospitals from taxation even if they provide no research, teaching, or subsidized care for indigents; that is, even if they are operated as strictly commercial nonprofits. Problems of contract failure do not seem important in the case of most hospital services. The continued predominance of the nonprofit form in this industry seems, instead, to be attributable to historical and financial factors largely unrelated to the relative efficiency of for-profit and nonprofit institutions. Moreover, there is evidence that, in general, the hospital industry is already overcapitalized. Thus, the hospital industry arguably fails both the criteria suggested above for administering the exemption. The current policy of exempting virtually all nonprofit hospitals may simply further encourage what already appears to be excessive capital investment in this sector.} \]

\[\text{Id.} \]
that for-profit hospitals simply do not provide.\textsuperscript{219} The special benefits of nonprofit, as compared to for-profit and government, hospitals include the provision of “more profitable services than government hospitals and more unprofitable services than for-profit hospitals.”\textsuperscript{220} Though Professor Horowitz does not conclude that these unique benefits of tax exempt nonprofit hospitals is caused by tax exemption,

\textsuperscript{219}\textit{See} Horowitz, \textit{supra} note 9, at 1347. Professor Horowitz explains:

The legal categories of corporate form matter a great deal. I present new empirical work showing that corporate form explains important differences in hospital behavior. I argue that not-for-profit firms very likely provide public and private goods that are both in the public interest, which for-profit firms fail to provide. By looking at only traditional measures of charitable behavior such as subsidized care for the poor, legal scholars have overlooked distinctions among ownership types. Instead, by examining the central function of hospitals - providing medical care - I find large differences among corporate forms, and argue that these imply large differences in hospital goals. Relying on this empirical work, I recommend that at least some hospitals in a market should be not-for-profit. We do not know enough to conclude which type of hospital or mix of types in a market is best. For the time being, we should assume that markets consisting of either entirely for-profit or government hospitals would not serve the public interest.

\textit{Id.}

\textsuperscript{220}\textit{Id.} at 1367. Professor Horowitz explains:

This part reports and interprets new evidence that comparable hospitals of different types - not-for-profit, for-profit, and government - offer different types of medical services. The findings imply that they implement different organizational goals. Although specifying these goals is difficult, the evidence supports the theory that government hospitals are hospitals of last resort. They are more likely than both other types to offer unprofitable services that are generally needed by poor, underinsured patients. For-profits seek profits and avoid offering unprofitable services more than the others. Not-for-profit hospitals are the intermediate type - while they are less responsive to financial incentives than are for-profits (both in offering profitable and avoiding unprofitable services), they are also less likely than similar government hospitals to offer unprofitable, undersupplied services. These results belie predictions that not-for-profit hospitals will behave no differently than for-profit hospitals in the production of public goods when under financial pressure.

\textit{Id.} at 1364.
she does acknowledge that this connection has not been disproven.\textsuperscript{221} Importantly, Professor Horowitz, consistent with this Article’s theory of contextual diversity, suggests that “[t]he near exclusive focus on charity care as an acceptable justification for tax exemption is too narrow. Tax policy should reflect the other important public benefits disproportionately provided by not-for-profit hospitals.”\textsuperscript{222}

\textbf{V. Potential Implications for the Structure of Tax Exempt Charity Law}

The analysis of the charitable tax exemption contained in this Article has several very important potential implications for the structure of tax-exempt charity law. In general, these implications center around the idea that the parameters of tax exempt charity law, though not endless, are at times unknown. While we can continuously re-evaluate what does and does not provide benefits to society and hence is entitled to tax exempt charitable status, we must be careful when proscribing particular functions as categorically non-deserving of charitable status. The reason for this hesitancy is that we simply never know what new and different value might be produced. A new perspective on an old activity may indeed be worthwhile. But we may never realize that value if we foreclose it categorically. This might be problematic, in some regard, for lawmakers and judges - for they necessarily have to draw lines and decide what is permissible and what is not. However, it is important to always recognize that those lines are not immoveable or static. Instead, the lines should be fluid and drawn from a variety of perspectives.

Similarly, the analysis in this Article also suggests that efficiency not be the only guide for how tax exempt charity law is crafted. Instead, we should draw on lessons from LMT theory that it takes many perspectives in order to obtain a clearer picture of the meaning ascribed to particular interpretations of the relationship among law, markets and culture. Thus, to the extent that we can draw on other than positive economic visions of law to make decisions about tax exempt charity law, we do so to our benefit. Accordingly, we should identify worthy and appropriate values for law and then think about the best way to approach and implement these values in

\textsuperscript{221}See \textit{id.} (“Whether the tax exemption causes the differences [] remains an open question. . . .[I]f the exemption is causing desirable not-for-profit behavior, then the costs of eliminating it may be high.”).

\textsuperscript{222}\textit{id} at 1349.
a market context. Using this line of reasoning, it is perfectly appropriate that CRT be used as a basis for gaining a better understanding of the public policy doctrine. The public policy doctrine emerged from a circumstance of racial discord. However, the policy adopted by the Supreme Court in *Bob Jones University* is devoid of racial components. Furthermore, no court since the Supreme Court in *Bob Jones University* has applied the public policy doctrine in any circumstance aside from racial discrimination. Given this state of affairs, contextual diversity would suggest that the public policy doctrine be invalidated and, instead, charities be explicitly prohibited from engaging in invidious racial discrimination. A nuanced interpretation of this particular conclusion would suggest that affirmative action, even if race-based, not be prohibited by this re-caste anti-discrimination rule.

**Conclusion**

Law involves a process of interpretation, and interpretation is socially situated. Tax law is no different. Although tax law is often represented by quantitative analysis in terms of its impact, this should not obscure the non-quantitative and interpretive aspects of tax law. Tax exempt charity law, and its allegiance to “mission” as opposed to “profit,” is a perfect vehicle for exploring the non-efficiency based aspects of tax law. This Article takes part in such analysis by articulating what is termed a “contextual diversity” theory of the charitable tax exemption. Contextual diversity requires that various aspects of the charitable tax exemption be examined, not only with the aim of maximizing efficiency, but also with the broader aim of advancing conceptions of justice that go beyond positive economic analysis to include fairness and other ideas important to a democratic society. Thus, in addition to using economic analysis to examine tax exempt charity law, scholars and others could possibly discover more diverse and different meanings in tax exempt charity law by drawing on appropriate non-economic legal approaches to law, such as CRT or others. This intellectual collaboration could not only broaden the discourse about the charitable tax exemption, it could potentially lead to discoveries about this area of law that we never knew existed. Thus, instead of thinking of the charitable tax exemption as simply an efficient means of providing certain goods and services to the public, our horizons might be broadened by thinking of the exemption in a different way. That is, we could think of the charitable tax exemption as a means of diversifying the market and, thus, allowing for more
creative and wealth producing opportunities. However, the charitable tax exemption is also subject to contextual constraints that act to limit the scope of charitable activity. In the end, the objective should be justice, not just efficiency.
Coping with Tax Competition in the EU, Preliminary Draft

Michael J. McIntyre
Professor of Law
Wayne State University

Preliminary draft, March 15, 2006. Please do not quote. I hope to include a revised version of this paper in a book on EU taxation that I’m writing with some Dutch colleagues.

1. Erosion of Power to Tax

In establishing the single market, the member states of the European Union (EU) have surrendered some of their sovereign control over their national borders. This loss of sovereign power has reduced their ability to imposes taxes of all sorts. The member states can recapture much of that lost power through various forms of concerted action. In particular, they could regain the ability to tax multinational corporations by harmonize corporate taxes within the EU and by prohibiting member states from engaging in predatory or beggar-thy-neighbor practices. The member states have harmonized their indirect taxes. No comparable action has yet been taken with respect to income taxes, although the loss of sovereign control over national borders poses a bigger threat to income taxes than to indirect taxes.

A tax is an involuntary payment; that is, taxpayers cannot be expected to pay a tax unless they are compelled to do so. A country that controls its borders can compel the payment of a tax by preventing those who do not pay from participating in the national economy. A member state of the EU, however, is required to allow residents of other member states, including multinational corporations, to exploit its market unconditionally.1 As a result, its ability to collect taxes, particularly corporate income taxes, is reduced significantly.

The reduction in sovereign power to tax resulting from free trade, free capital movement, and the associated four freedoms provided under the various EU treaties has increased the vulnerability of the member states of the EU to what is commonly referred to as “tax competition.” That term is used to describe the actions taken by a country to assist the taxpayers of another country in...

1Free trade agreements generally have this effect, although most free trade agreements are not as comprehensive as the open-border agreements within the EU.
evading or avoiding the taxes imposed by another country by offering those taxpayers a more favorable tax regime. In many cases, the tax regime offered is a complete exemption from income taxes.

By providing the taxpayers of a country with the ability to evade or avoid that country’s tax, a low-tax country engaging in tax competition may create a situation in which the victimized country feels compelled to reduce its own tax rate. This pressure to lower tax rates is sometimes referred to as a “race to the bottom.” The point made by that term is that countries competing with each other to have the lowest tax rate will eventually end up with a tax rate of zero. They may well have a negative rate on income from marginal investment.

The member states of the EU have responded to tax competition from countries both within and without the Union by lowering their effective corporate tax rates, in a few cases to zero. Some member states have reduced their statutory tax rate, whereas others have offered various tax incentives for new investment. Countries offering tax incentives typically are attempting to exempt corporations on the income they derive from new investments while continuing to tax them on income derived from existing investments.

Countries engaging in tax competition often have no significant economic relationship either to the taxpayers who are avoiding tax or to the income of those taxpayers. They do not actually attract real foreign investment but rather allow the shifting of income from the victimized state through various paper transactions.

In some cases, however, the taxpayer of the victimized country may actually shift some productive activities to the low-tax country. In this latter case, the country from which the production activity is shifted is a double loser, losing both the tax revenue and the economic benefits it would have obtained from having the production activity take place within its own borders.

Many countries, including several member states of the EU, have taken unilateral action to cope with what they believe to be tax-haven abuses involving merely paper transactions. These measures, which have been moderately successful in some cases and could be more successful if pursued with greater skill and vigor, are generally ineffective in dealing with tax avoidance involving low-tax countries within the EU. One reason they are ineffective is that decisions of the European Court of Justice generally prevent a member state from taxing outflows to another member state, even if that other member state is operating as a tax haven.

2. Nature of Tax Competition

The term “tax competition” is a misnomer. A low-tax country that assists the residents of other states in evading or avoiding taxes otherwise due is not
“competing” with the victimized country, any more than the lion is “competing” with the wildebeest when it seeks to have the wildebeest for dinner. The low-tax country seeks to advance its national agenda by dividing the spoils of tax evasion or avoidance with taxpayers who otherwise would have paid tax in the victimized country. The victimized country, in contrast, is merely seeking to preserve its sovereign right to set its own tax policies; it is not attempting to extract any benefit from the low-tax country.

Some apologists for tax competition have suggested that tax competition prevents governments from “overcharging” taxpayers through inappropriately high taxes in much the same way that competition among private producers of goods and services prevents them from overcharging their customers. This analogy of tax competition to real competition is more likely to mislead than to enlighten. Real competition prevents producers from artificially reducing the supply of goods and services in order to drive up prices. That is, real competition is the antidote to private monopoly. Tax competition has no comparable effect. The idea that governments are seeking to drive up the “price” of governmental services by reducing their supply is bizarre. To the contrary, if tax competition has any effect on the supply of government services, it is to reduce that supply rather than to increase it.

Unlike private entrepreneurs, governments generally seek to deliver their services only within their national borders. They do not attempt to export those services to other countries. France does not aspire to build an Autobahn in Switzerland or a dike in the Netherlands, and Poland is not interested in delivering the mail in Paris or in redistributing income among residents of Spain. As a result, each government has a nearly exclusive “market” for the services it delivers. Tax competition does not change this fundamental feature of government, and, as a result, it has no effect on the price charged for government services or on the quality of those services.

Although taxes may be the price paid for civilization, they are not the price paid by individual taxpayers to obtain a particular quantity of government services. Income taxes are used by national governments primarily to pay for public goods, such as public works and public order, and to redistribute income in accordance with ability to pay. The relationship between the amount that a particular taxpayer pays in taxes has no systematic relationship to that taxpayer’s access to public goods. Indeed, some taxpayers who pay the least in taxes obtain the largest amount of public goods. In addition, the benefits that a taxpayer obtains from a government’s redistribution of income typically relate inversely to the size of the taxpayer’s tax bill. In brief, income taxes generally are not a market mechanism for allocating control over government resources among taxpayers.

Tax competition can reduce a government’s tax revenues, but the most likely effect in most European countries is to change the tax mix, causing member
states to rely more heavily on the VAT, which is a regressive tax, and less heavily on progressive income taxes. That result will be viewed as undesirable by those favoring progressive taxation and as desirable by those favoring regressive taxation. The result is unambiguously bad in one respect — it prevents the people in each of the countries of Europe from deciding for themselves, through the democratic process, the type of tax system they would like to have.

Some economists have asserted that tax competition in Europe might have efficiency gains by allowing taxpayers to vote with their feet, so to speak, for the type of government that they prefer. The theory, known as the Tiebolt hypothesis, is that if some member states offered their residents high taxes and a high level of government services and other member states offered low taxes and a low level of government services, then taxpayers could maximize their welfare by moving to the country that had the tax-benefit mix they preferred. In addition, by watching the migration patterns, a government allegedly would obtain useful information about the preferences of its taxpayers for particular public goods.

One of the assumptions of the mass-migration theory described above is that migration from one community to another is a benign event. Mass migration within Europe, however, certainly would not be viewed as benign by many Europeans. They are more likely to view mass migration as an unmitigated political disaster — a serious threat to the cohesion of the Union and to the distinctive national characteristics they are hoping to preserve within that Union.

Another implicit assumption of the mass-migration theory is that governments are using tax revenues solely to finance government services. Within Europe, that assumption is clearly false. Most member states spend a substantial part of their tax revenues to reduce the inequalities in the distribution of income and wealth that typify a market economy. In addition, these countries have adopted progressive income taxes to mitigate income inequalities. As noted above, tax competition is a threat to those redistributive policies. In any event, migration of some wealthy taxpayers from high-tax countries to low-tax countries would at most confirm that those wealthy taxpayers are not keen to pay taxes according to their ability to pay. It would reveal nothing significant about their preferences for public goods.

Notwithstanding the claims of its supporters, the mass-migration theory cannot justify tax competition in Europe even if the dubious assumptions underlying the theory were accepted as valid. The theory was developed to apply to small, homogeneous communities with only one type of tax. Its application to national governments with large and diverse populations, a multiplicity of diverse spending programs, a multiplicity of taxes, and a commitment to redistribution is

---

questionable at best. Even if the theory could be applied to nation states, it would not justify tax competition in Europe. At most, it would justify some diversity in the mix of taxes and benefits within Europe — a result that already exists and is more likely to be preserved through cooperative efforts than through a race to the bottom.

Some commentators assert that tax competition has beneficial distributional effects by shifting productive activities from highly developed countries to less developed countries. A rich country that acts though its political processes to share its abundance with poor countries deserves high praise. The likelihood that tax competition can serve as an effective instrument for shifting substantial production from rich countries to poor countries, however, is remote in the extreme. What some commentators have perceived to be a desirable effect of tax competition is really little more than the manipulation of defective international tax rules by multinational companies to evade or avoid taxation in the country where the income actually arose.

For the poorest of the developing countries, the main effect of tax competition has been to reduce their already inadequate ability to mobilize resources for development. With few exceptions, they have not experienced a significant increase in foreign direct investment over the past decade, notwithstanding the huge increase in tax competition and the worldwide increase in investment capital. Most foreign direct investment goes to the developed countries, and the share going to the developing countries goes primarily to those toward the high end of the development spectrum.

A poor country has great difficulty attracting productive foreign investment through income tax concessions because poor countries generally do not offer many opportunities to potential foreign investors to earn substantial profits from business activities conducted within their borders. As a practical matter, the anticipated profits in the poor country, discounted for risk, would have to exceed 70 percent or more of the before-tax profits anticipated in high-tax countries for an income tax exemption to be attractive. A poor country might be able to offer a few such investment opportunities, but a country able to offer a significant number of them almost certainly would not be a poor country.

The kinds of productive activities that a poor country is most likely to attract through income tax incentives are activities that are relatively mobile — that is, activities that could be conducted in many different locations. A common example of a mobile production activity is contract manufacturing. Such activities typically generate profit margins close to the average rate of return on capital because many competing companies could conduct those same activities and would do so if the profit margins were attractive to them.

Although a few poor countries have had success in using tax incentives to attract productive investment, the reason for that success is not generally
understood. As noted above, most multinational companies will not be induced to move productive activities to a poor country, given all the risks typically associated with such investment, merely to shelter from taxation the low profits those activities are likely to generate. What they do find attractive is the ability to use the productive activity in a low-tax country to shelter profits arising from activities conducted elsewhere. As noted above, the unilateral anti-avoidance rules adopted by some countries have been moderately effective in limiting the use of tax havens when only paper transactions are involved. Those rules generally are ineffective, however, when productive activities are actually conducted in the low-tax country. As a result, multinational companies that may appear to be seeking a low-tax environment for productive activities are actually seeking to use essentially paper transactions, combined with limited production activities, to shelter income derived in other countries.3

The basic pattern described above is illustrated by the Bosch & Lomb case, involving a tax dispute between a contact lens company and the US tax authorities. The contact lens manufacturing business, initially conducted almost exclusively in the United States, was hugely profitable. The taxpayer opened an assembly and fabrication plant in Ireland to produce the contact lenses there. The operations in Ireland were quite routine and normally would not have been expected to generate above average profits in Ireland. The US tax authorities, supported by expert testimony, characterized the operations in Ireland as a form of contract manufacturing, available in the market for some modest return on invested capital. The taxpayer was successful in that case in allocating very substantial profits to Ireland, due to perceived imperfections in the US tax rules. It was that tax benefit — avoiding tax on income generated outside Ireland — that made the Irish tax concessions particularly valuable to the foreign investor.

Whatever the general merits of the argument for tax competition on distributional grounds, it cannot justify tax competition within the European Union. Tax competition is a wasteful method of achieving redistribution among countries because most of its benefits go to multinational companies, with only at most a small benefit trickling down to the poor countries. The EU has no reason to use that crude and possibly ineffective method when it has available a fully effective method — direct grants from the EU to those EU countries with a below-average standard of living. This grant method is described above in Chapter 1.B. An added advantage of a direct grant program is that it has been adopted in accordance with democratic procedures and can be adjusted by the member states to take account of changed circumstances.

3. Harmful Tax Competition

From the perspective of the victimized countries, tax competition is always harmful. These countries lose the tax revenues they otherwise would have obtained and gain nothing in return. In addition, the aggregate impact of tax competition on government revenues is always negative. That is, the revenue losses suffered by the victimized countries always exceed the benefits obtained by the low-tax countries. Obviously, taxpayers would not voluntarily shift income from one country to another for tax reasons unless the effect was to reduce their overall tax burdens.

From the perspective of the multinational companies and other taxpayers who have their tax bills reduced, tax competition is never harmful. On the contrary, it is a bonanza. It is not surprising, therefore, that the multinational companies are enthusiastic promoters of tax competition in all of the countries in which they have influence.

Tax competition is helpful to some of the countries that are able to share with foreign taxpayers the tax savings resulting from tax competition. As discussed above, that benefit comes primarily from exploiting weaknesses in the international tax laws of foreign countries, not from any inherent value to tax competition.

Many of the practitioners of tax competition are themselves victims of the economic forces they seek to exploit. In particular, they often are compelled by market pressures to reduce or eliminate corporate income taxes on their own taxpayers to avoid giving an unfair and politically unacceptable competitive advantage to foreign-based multinational companies. In addition, the benefits that some countries actually obtain from practicing tax competition are modest in comparison to the benefits they would obtain from the creation of an international framework in which countries were able to impose moderate taxes on corporations.

In general, a country is unlikely to benefit from tax competition if it is so poor that it does not offer foreign investors many opportunities to earn a significant return on their investment. A country is also unlikely to benefit from tax competition if it is so far along the road to development that it would relinquish significant tax revenues from its domestic taxpayers by offering a low-tax regime.

The Council of Europe and the OECD have proposed that the world in general and the EU in particular take practical steps to eliminate certain types of tax competition. They have characterized the tax practices they hope to eliminate as “harmful tax competition.” That phrase is unfortunate, for it seems to imply that some tax competition is beneficial, or at least benign. As noted above, however, all tax competition is harmful to the victimized country and has an
overall negative effect on government revenues, economic efficiency, and distributive justice.

What the Council of Europe and the OECD actually have been trying to identify are tax practices that can be considered as illegitimate exercises of sovereign power. In these cases, it is thought to be appropriate for the victimized governments to take unilateral or collective action to persuade or pressure the offending governments to change or abandon those practices.

No government can be fairly criticized simply because it has decided to exempt from tax the income derived within its borders, or to tax that income at a low rate. The power to tax is an important power of a government, and that power implies the power not to tax if the sovereign so chooses. The fact that a failure to tax causes harm to some other countries does not give those countries a legitimate cause for complaint.

A government can be fairly criticized, however, if it conspires with foreign taxpayers to undermine the power of other governments to tax income arising within their own borders. In addition, a government can be fairly criticized for taking advantage of the inevitable weaknesses in the international tax rules of other countries to poach on their tax base. Although the line between acceptable and offensive conduct may not always be clear, that line is clear enough in many important cases to justify collective or unilateral action to combat the offensive conduct.