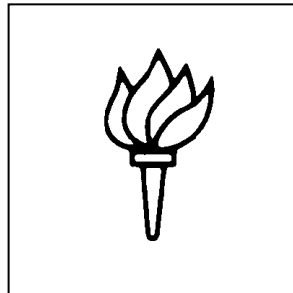


NEW YORK UNIVERSITY

SCHOOL OF LAW

PUBLIC LAW & LEGAL THEORY RESEARCH PAPER SERIES
WORKING PAPER NO. 14-04

LAW & ECONOMICS RESEARCH PAPER SERIES
WORKING PAPER NO. 14-06



The Economics of Tax Law

Daniel Shaviro

February 2014

THE ECONOMICS OF TAX LAW

Daniel Shaviro*

Revised Draft

All Rights Reserved

January 2014

1 INTRODUCTION

Many volumes could be, and have been, devoted to the question of how the law of, say, torts, contracts, corporations, civil procedure, or intellectual property could be designed with an eye to maximizing economic efficiency. In tax law – leaving aside Pigovian taxes, which address negative externalities – the question of how to maximize efficiency is almost childishly easy to answer. Unfortunately, however, the answer has very little direct relevance to understanding either existing law or potentially desirable tax reform.

Suppose the government has a fixed cost of providing public goods, and that tax revenues will be used to meet this cost. Or suppose instead that whatever tax revenues are collected will be handed out again to various people, in amounts that are unrelated to how much tax they individually paid. In either case, taxation involves an externality problem. When people make choices that would affect their tax liabilities, they only have reason to care about their after-tax positions. But the tax payments go to someone,¹ creating a positive revenue externality.

Thus, suppose there is a 30 percent wage tax and that, by working for an hour, I could earn \$100 before tax. In deciding whether to take the job, I will only consider the \$70 that I would get to keep, not the full \$100 pretax wage that presumably reflects the value of my

* Wayne Perry Professor of Taxation, NYU Law School. I am grateful to Alan Auerbach and Mitchell Kane for their comments on an earlier draft. This paper is a work in progress and is forthcoming in print in *The Oxford Handbook of Law and Economics*, edited by Francesco Parisi, and online at <http://www.oxfordhandbooks.com/>.

¹ Tax revenues that pay for providing public goods can be thought of as going to whomever would otherwise have borne the cost of such provision.

production to someone else. If my reservation wage, reflecting my disutility of working, is between \$70 and \$100 and I thus turn down the job due to the tax, then not only is no tax revenue generated in this instance, but the wage tax prevents the creation of surplus (from the excess of \$100 over my reservation wage). It thereby gives rise to deadweight loss.

Given revenue externalities and the consequent prospect of deadweight loss, maximizing the efficiency of non-Pigovian taxes is straightforward. They should be lump-sum – that is, invariant to taxpayer decisions. (Or they might be effectively lump-sum, by reason of depending on decisional margins that are wholly inelastic.) If I will owe the government the same amount of tax no matter what I do, tax-induced deadweight loss will not arise

The best-known example of a lump-sum tax is a uniform head tax – assuming that one truly cannot avoid it, such as by exiting the jurisdiction or being too poor to pay. Likewise, consider a height tax, if people cannot affect their height (see Mankiw and Weinzierl 2010). As we will see below, the optimal income tax (OIT) literature, founded by Mirrlees (1971), posits the theoretical possibility of a lump-sum tax based on “ability,” taken to be an entirely fixed trait, but assumed not to be directly observable, making it unavailable as a tax instrument.

From the standpoint of efficiency, no lump-sum tax is better than any other – by definition all succeed equally in avoiding the creation of deadweight loss.² This leads directly to two main questions. First, why are lump-sum taxes, or instruments that come as close to them as possible, so absent, not just in actual practice but even in theoretical debate about tax policy? The answer turns on the importance of distributional issues. Second, how do considerations of efficiency operate once we have accepted, for distributional reasons, the need for tax instruments

² Lump-sum taxes may affect behavior, without creating deadweight loss, by reason of income effects. For example, work decisions may be influenced by lump-sum taxes that affect the resources one has available to meet basic needs. My reservation wage might be higher, for example, when lump-sum taxes are high than when they are low. However, lump-sum taxes will not induce me to reject jobs that offer more than my reservation wage, or to accept those that offer less.

that (like the wage tax note above) have the unfortunate side-effect of discouraging productive activity?

2 DISTRIBUTIONAL ISSUES

A longstanding tradition in tax policy thinking holds that tax burdens should vary based on “ability” or “ability to pay” (Murphy and Nagel 2002, 20). Under this view, Bill Gates should pay more tax than a middle class worker, who should pay more tax than a homeless person. The question of whether this involves “redistribution” from high earners to low earners depends on one’s baseline for assessing the “re” in “redistribution.”³ However, subject to questions of tax incidence, a system that does this clearly is progressive relative to levying a uniform head tax, which presumably would be acceptable in the absence of distributional concerns.⁴

There are numerous potential motivations for seeking greater after-tax distributional equality than that which would result from levying a uniform head tax. For example, one’s social welfare function might treat greater equality as good in itself. Or one might believe that inequality yields negative externalities, such as from its aiding capture of the political system by the rich that leads to pervasive rent-seeking. Or one might view it as limiting economic opportunity for children born into the lower strata, or as yielding a more hierarchical and stratified society that generates net unhappiness.

In a conventional welfare economics framework, the chief rationale for requiring high-earners to pay more tax than low-earners is the paired set of assumptions holding that people’s utility functions (a) are similar and (b) feature declining marginal utility (DMU) for wealth. Of

³ Redistribution by the tax system implies that there is a pre-tax or nontax baseline in which wealth was differently distributed, but what that means, and how it should be interpreted, is unclear.

⁴ The aim of pursuing distributional aims through the tax system rests on the views, not only that distributional concerns are relevant, but that the tax system is a suitable place to address them. In the law and economics literature, this is generally accepted, with the main point in contention being whether other parts of the legal system should *also* play a role. See, e.g., Kaplow and Shavell (1994); Sanchirico (2006).

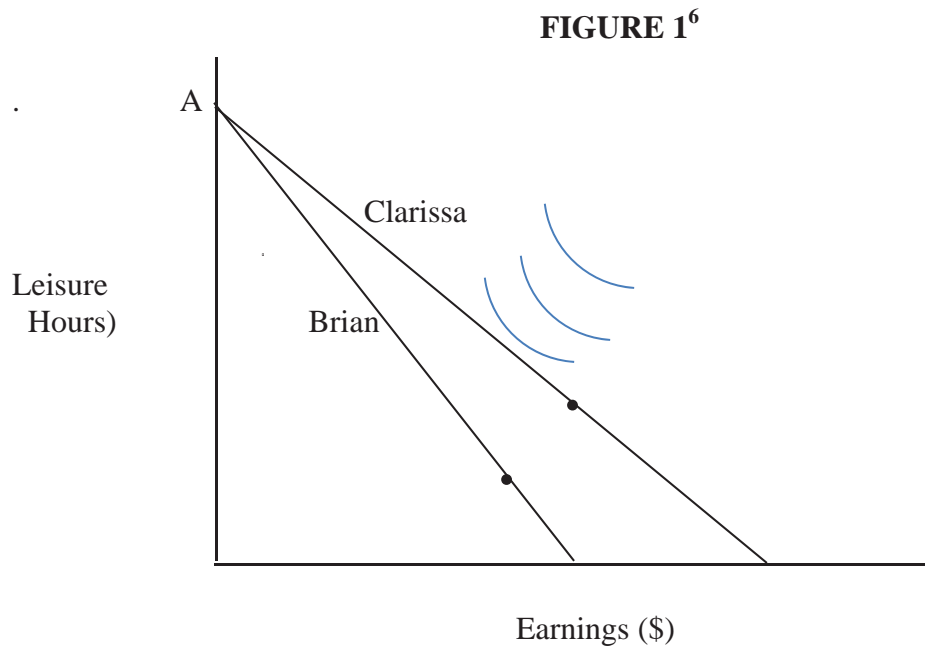
these two assumptions, while (a) might be viewed in the main as a non-falsifiable default stance, (b) arguably is supported empirically by the prevalence of risk aversion, and explained theoretically on the ground that we generally satisfy our most urgent needs first, moving to the ever more psychically dispensable as our budget lines move outward. Acceptance of the two above assumptions suggests that, all else equal, the disutility from bearing a dollar of tax burden steadily declines as one's material circumstances improve.

Suppose that one wholly ignores incentive effects, assumes that people are relevantly identical other than in how well-off they are, and treats DMU as the sole relevant consideration in assessing their welfare losses from bearing tax burdens. The conclusion that would follow, under a welfarist approach to tax policy,⁵ is that the government should completely equalize after-tax distribution, such as by expropriating all wealth (i.e., taxing it at a 100 percent rate) and then distributing it pro rata (Edgeworth 1897, 553). However, ignoring incentive effects is no more plausible, from a welfare standpoint, than ignoring distributional effects and thus endorsing a uniform head tax. The central task that both a 100 percent tax and a uniform head tax ignore is how to handle the tradeoff between efficiency and distributional concerns.

The OIT literature squarely addresses this tradeoff. A very basic model, derived from Mirrlees (1971), might proceed as follows. Suppose that people derive utility from market consumption and from leisure, each of which has DMU. Market consumption is funded through work (this being a one-period model, and hence without saving). People differ only in ability or wage rate, defined as the amount that one can earn per unit of time or effort – inputs that are not themselves directly observable. Ability is innate, but cannot directly be observed. By contrast, earnings, which are the joint product of ability and time or effort, can be observed. People can

⁵ A welfarist or welfare economics approach can be defined as one in which “social welfare is posited to be an increasing function of individuals’ wellbeing and to depend on no other factors” (Kaplow and Shavell 2002, 24).

work as much or as little as they like, up to a maximum (such as the number of waking hours in a day). Each individual chooses the level of labor supply that would place her on the highest indifference curve that is tangent to the budget line resulting from the available time and the wage rate – as in Figure 1, which features two individuals with different wage rates.



Not surprisingly, in light of their identical utility functions, Clarissa, who has the higher wage rate of the two, ends up with both more earnings (and thus more market consumption) than Brian, and more leisure. Given their identical utility functions, featuring DMU for market consumption (as well as leisure), taking money from Clarissa and giving it to Brian would, up to a point, increase his utility more than it would reduce hers. The mechanism employed is a tax on earnings that funds a uniform cash grant – resulting in redistribution since Clarissa pays more tax than the amount she gets back, while Brian pays less.

The income effect of this transfer, considered in isolation, would induce Brian to work less than if the tax-plus-grant program did not exist, and Clarissa to work more.⁷ This, however,

⁶ Figure 1 is adapted from Shaviro 2007, 755.

merely reflects re-optimizing by both parties.⁸ More problematic is the program's substitution effect, inducing both parties to work less, given that they would lose a portion of their earnings to the tax collector. The program therefore discourages labor supply, creating deadweight loss and reducing the amount of redistribution that is desirable.

Unsurprisingly, the OIT literature finds that tax rates and the size of the grant are negatively correlated with labor supply responsiveness. They are positively correlated with (1) the slope of the DMU curve (and thus the welfare gain from each dollar that is redistributed), (2) the degree of dispersion in ability levels, and (3) any pro-egalitarian weighting in one's social welfare function (Slemrod 1990, 165). They also should be positively correlated with the size of any negative externalities that result from wealth inequality.

An apparent key finding of the OIT literature, accepted for many decades though more recently challenged, was that tax rates should be relatively flat (Slemrod 1990, 165). The underlying reasoning is straightforward. On the one hand, DMU supports rate graduation, because the last dollar earned by a high earner purchases less utility than the last dollar earned by a low earner. On the other hand, however, lower rate brackets are effectively inframarginal for those who know with certainty that they will earn enough to place out of these brackets.⁹ This causes such brackets generally to offer more revenue bang for the buck, in efficiency terms, than higher rate brackets, which tend to be inframarginal for a smaller percentage of the individuals who are subject to them. This consideration pushes back against rate graduation (Shaviro 2011,

⁷ Having more wealth reduces the marginal utility of the last dollar earned, so Brian would be expected to shift towards greater consumption of leisure. By contrast, having less wealth induces Clarissa to re-optimize by working more.

⁸ It is true, however, that the income effect on Brian has a negative revenue externality, while that on Clarissa has a positive revenue externality.

⁹ Thus, consider the Social Security portion of the payroll tax, which imposes an aggregate 12.4 percent tax rate on wages up to its dollar cap (\$117,000 for 2014). This tax is projected to raise more than \$600 billion annually, yet it has no effect on the marginal work incentives of people who are certain to earn more than the taxable amount.

832), and indeed offers theoretical support for the conclusion that the marginal rate at the very top of the income distribution should be zero.¹⁰

In many OIT models, these offsetting considerations have yielded a marginal rate structure that is relatively flat. However, Diamond and Saez (2011) argue, within the OIT framework, for steep rate graduation at the top. Suppose that “the social marginal value of consumption for top-bracket tax filers” is so low that “as a first approximation [it] can be ignored” (168) – i.e., treated as indistinguishable from zero.¹¹ Then the socially optimal marginal tax rate for such high-end income is the revenue-maximizing rate, which Diamond and Saez argue may exceed 70 percent (171).¹² What is more, if one assigns close-to-zero weight to marginal welfare losses incurred by people at the top of the income distribution, “there is little lost if we adopt a policy of increased marginal tax rates that ends up losing a small amount of revenue” (Auerbach 2012, 703). If the exact revenue-maximizing rate at the high end is uncertain, one might even view as symmetric the social cost of mistakenly setting the top rate below the revenue-maximizing level, and above it.

3 EFFICIENCY ISSUES

Once we are taxing productive economic activity for distributional reasons, rather than imposing a lump-sum levy of any kind, the aim of efficiency analysis of taxation is to minimize the harm done, rather than to achieve a first-best optimum. Here two alternative approaches – directly inconsistent with each other – may apply. The first is neutrality, while the second is

¹⁰ As Slemrod (1990, 164) notes, “raising the marginal tax at the [very] top above zero distorts the labor supply decision of the highest earner but raises no revenue.” However, because this result only holds “precisely at the top,” he concludes that it “should not be taken seriously as a practical guide to tax policy.”

¹¹ This might reflect additional consumption’s lack of marginal utility to the high-earner herself, who might already have far more wealth than she could ever expect to spend. Or it might reflect egalitarian social weighting, or the view that any positive marginal utility to consumption at the top of the income scale is likely to be offset by negative externalities associated with high-end wealth concentration.

¹² The assumptions underlying this result include the existence of a fat tail at the top of the income distribution, and the absence of a finite maximum income level.

differential taxation in response to any of several distinct rationales that include varying elasticities.

3.1 Tax Neutrality

In a simple one-period OIT model, the tax base is earnings, or equivalently market consumption (since savings would require adding the dimension of time). Once we make things more complicated, however, tax base design becomes an issue. Even leaving aside the tax treatment of savings (discussed below in relation to the choice between income taxation and consumption taxation), there is the question of whether all items that are potentially in the base – for example, all earnings or all consumer commodities – should be treated the same by the tax system.

The efficiency case for tax neutrality, as between types of work and/or alternative commodities, is straightforward. Given the point that paying taxes gives rise to a positive revenue externality, deadweight loss results when taxpayers switch from higher-taxed to lower-taxed choices by reason of the tax disparity. Only through tax neutrality, in the sense of taxing equally valuable items the same, can one avoid tax distortions with respect to the choice between taxable items. Moreover, even the fact that one is generally discouraging labor supply and market consumption via the tax does not change this analysis, absent further considerations that are discussed below.

This analysis generally holds even where tax disparities would have progressive distributional effects. Thus suppose one is considering imposing an extra tax on luxury goods, such as yachts, where the alternative would be to raise tax rates on the rich. At a first approximation, the luxury tax does nothing to ease the adverse incentive effect on labor supply, since high-earners will still face a higher tax burden if they work more in order to finance a yacht

purchase. In addition, however, one is distorting their consumption choices between yachts and other items. A similar analysis applies to tax-favoring necessities (which may form a higher percentage of low-earners' than high-earners' consumption bundles) for distributional reasons.

3.2 Deliberately non-neutral taxation

The efficiency case for tax neutrality within the realm of productive economic activity and market consumption assumes the inapplicability of a number of considerations that might require modifying the analysis. For example, choices that have positive or negative externalities should be encouraged or discouraged, respectively, whether or not the Pigovian element is administratively integrated with the distributionally minded tax instrument. Moreover, if people have cognitive difficulties in optimizing their pre-tax choices, then, whether or not tax incentives and disincentives (as distinct from “nudges”) can be counted on to improve things, the actual payoff to neutrality may be reduced or put into question. Also, taxing particular choices that serve as a signal or tag of ability or wage rate can potentially reduce overall distortion even though it inefficiently discourages those particular choices.¹³

There is also, however, a potential rationale for differential taxation that responds to the basic set-up of the optimal tax problem – that is, the fact that only earnings and/or market consumption generally fall within the scope of the tax. Consider the optimal commodity tax literature (derived from Ramsey 1927), which addresses efficient tax rate-setting for different consumer goods, and finds that varying rather than uniform rates may be optimal. The basic setup here is as follows. Suppose the government is trying to raise revenue through taxes on specific consumer items or commodities. What is more, suppose that commodity taxation's sole

¹³ Suppose, for example, that one reduced high-end marginal tax rates while imposing a tax burden of some kind on the decision to attend elite universities. If attending such universities is a signal or tag that is associated with high ability (as distinct from being a way to increase one's earning potential), then it is theoretically possible that this would permit reducing the efficiency cost of redistribution.

purpose is to raise revenue for the government at the lowest possible efficiency cost – rather than, say, either to address externalities like a pollution tax, or to address distributional goals. Finally, suppose that lump-sum taxes, such as uniform head taxes, are unavailable, as are comprehensive income or consumption taxes.

Against this background, one might initially suppose that it would be most efficient for all of the taxable commodities to face the same tax rate. Ramsey showed, however, that the goal of minimizing overall deadweight loss, by equalizing each particular commodity tax's marginal deadweight loss, had instead a very different implication. In particular, suppose that some commodities are more tax-elastic than others, meaning that taxpayers would substitute away from them more if all commodities were taxed at the same rate. In general, the deadweight loss per dollar of revenue raised generally rises with the tax rate. Thus, each commodity tax would be expected to feature a rising ratio of deadweight loss to revenue as one increases its rate. Under these conditions, the way to minimize deadweight loss, while keeping overall revenue constant, is to have a higher tax rate on commodities that are relatively inelastic, and a lower one on those that are relatively tax-elastic (Gruber 2011, 602-603).

With an OIT-style income tax in place, the relevance of arguments drawn from optimal commodity taxation narrows significantly. When one adds distributional considerations to the analysis and assumes, unlike Ramsey, that one *can* uniformly tax all consumer goods, uniformity is generally best after all (Atkinson and Stiglitz, 1976). The key underlying idea here is that one needs to embed commodity choice in a broader framework, in which people are deciding how much to work and earn. After all, without earnings one cannot pay for consumer goods. In this framework, it turns out that the only reason for differentially taxing particular commodities relates to the underlying labor supply choice that both the income tax and more limited

commodity taxes distort, by discouraging market consumption (and therefore work). In light of this efficiency problem, differential commodity taxation can reduce the labor-leisure distortion if it involves tax-favoring leisure substitutes – that is, commodities that tend to be consumed more when leisure declines – and/or tax-penalizing leisure complements – that is, goods that tend to be consumed jointly with leisure (Corlett and Hague 1953).

Suppose, for example, that restaurant and take-out meals are leisure substitutes, due to their appeal to people who are working long hours, while groceries are a leisure complement, because one needs free time to cook at home. Then it may be efficient to tax groceries and subsidize restaurant meals, as a way of lowering the labor-leisure distortion (Kaplow 2008, 138 n. 16). Otherwise, however, “results derived in the original Ramsey framework, in which no income tax is available, fail to provide proper guidance in a world with an income tax” (148).¹⁴

3.3 Political economy

As the previous discussion shows, despite the revenue externality problem that pushes in favor of tax neutrality, there are a number of reasons why well-tailored departures therefrom may increase efficiency without undermining the achievement of distributional aims. There is also, however, a political economy argument for generally treating tax neutrality as a core principle. Suppose that, in political and administrative practice, tax neutrality is easier to evaluate and apply than the variegated set of arguments that may support treating particular items more favorably than others. Then, in a political environment rife with agency costs, interest group influence, and low information on the part of voters, generally steering for greater rather than lesser tax neutrality may be likely to result in the achievement of greater tax efficiency in practice.

¹⁴ But see Shaviro 2014 (164-165) for an example in which the conclusion from Kaplow 2008 does not apply because the requisite income tax is not sufficiently available (e.g., because resident individuals cannot be taxed on foreign source income earned through foreign corporations).

3.4 Relevance of administrative constraints

The discussion thus far has mainly been at a high level of abstraction. For example, it generally assumes the feasibility of a comprehensive tax on people's earnings and/or on all market commodities (though not leisure), with behavioral responses being limited to those that involve labor supply or commodity choice – as distinct from, say, tax evasion or avoidance. For both political and administrative reasons, actual income taxes are generally quite far from having so simple and straightforward a character.

Some departures from the idealized model involve preferential treatment for particular items or activities – for example, in the United States, owner-occupied housing and employer-provided health insurance. Often, such tax preferences are widely viewed by tax policy experts as having little merit. Even where arguably meritorious, they may have design problems that reflect their having taken the form of “tax expenditures” – that is, spending or allocative policy that is funneled through the tax system – in large part for optical reasons, so that officially measured “taxes” and “spending” will appear to be lower (see Shaviro 2004a).

Real-world income tax systems also commonly reflect administrative constraints that can cause them to depart greatly in practice from the idealized models. Examples include the realization requirement, under which gains and losses from changes in asset value are not accounted for by the tax system until a transaction such as a sale occurs, and the primary use of entity-level, rather than owner-level, taxation of corporate income. These rules can cause huge swathes of economic income to be taxed very differently than close economic substitutes, and can also induce substantial deadweight loss within their operating realms.

Further, behavioral responses to taxation can involve, not just labor supply and commodity choice, but a spectrum of tax avoidance and evasion behaviors, in which legal

formalities and/or enforcement efforts play paramount roles in determining what is paid. The issues thereby raised greatly complicate optimal tax design and our ability to understand the practical consequences of a given set of tax rules (Slemrod and Yitzhaki 2002).

These features of real-world tax systems create complicated second-best problems for policymakers. For example, they may support the enactment of rules that deliberately mismeasure the taxpayer's income in one respect, as an offset to what is at least suspected mismeasurement in other respects. Two illustrations, drawn from the U.S. federal income tax system, are the following:

(a) Indirect partial repeal of the exclusion for imputed rental income: Homeowners generally can exclude the imputed rental income that they derive by living rent-free on their own properties. The exclusion distorts choices between home ownership and rental, as well as between home and non-home consumption. If we take its existence as given for political or administrative reasons, but wish to address its distortionary effects, one way to proceed would be by denying deductions for home mortgage interest.

Suppose a homeowner has a mortgage loan, generating annual interest expense. Interest deductions generally should be allowed in a comprehensive, well-functioning income tax, as they are (negative) returns to (dis-)saving, and thus the mirror image of positive returns to saving that an income tax reaches. However, denying home mortgage interest deductions can function as the equivalent of partial, indirect repeal of the exclusion for imputed rental income (Shaviro 2000, 144). In fact, the U.S. rules allow home mortgage interest deductions, subject to certain limitations,¹⁵ but this is widely viewed as a deliberate tax preference that, while strongly entrenched politically, attracts little normative support in academic tax policy debate.

¹⁵ Home mortgage interest is allowable only as an itemized deduction, thus denying it to some individuals who claim the standard deduction in lieu of itemized deductions. In addition, the mortgage loan principal that can generate

(b) Loss limits and nonrefundability: In a comprehensive, well-functioning system, just as economic gains would be fully taxed, so economic losses would be fully deductible. This would include the provision of overall refundability (i.e., the allowance of net refunds). If one instead has loss limits and nonrefundability, one discourages risk-taking, by causing investments with greater variance to face higher expected tax burdens than those with the same expected pretax returns but less variance.¹⁶

Nonetheless, the U.S. tax system not only makes overall losses nonrefundable (although they may be allowed against net taxable income in other years), but disallows net losses in particular income categories. For example, net capital losses generally cannot be deducted in a given year.¹⁷ The rationale for such rules is that, given tax planning opportunities, one cannot be confident that taxpayers claiming losses have actually suffered a net loss economically. For example, despite a wide range of anti-tax shelter rules, taxpayers may be able to arrange transactions in which, due to the realization requirement and the heterogeneity of tax accounting rules, circular cash flows generate enormous claimed losses. To explain this generically, suppose that I both pay and receive \$1 billion from the same counterparty (which can be accomplished simply by swapping pieces of paper). If I am able to claim, for some technical reason, that the \$1 billion I paid is currently deductible, while the \$1 billion that I received is not currently includable, then I have a net \$1 billion loss deduction. This amount can easily be

allowable interest deductions is capped at \$1 million for acquisition indebtedness, plus an additional \$100,000 of other indebtedness that is secured by the home. See U.S. Internal Revenue Code section 163(h)(3). These amounts are not indexed to inflation, and thus may gradually decline in real terms.

¹⁶ Graduated marginal rates generally have this effect, reflecting that nonrefundability can be viewed as an example of rate graduation (i.e., from zero to a positive rate as net income crosses the zero threshold).

¹⁷ See U.S. Internal Revenue Code section 1211. Individuals can deduct net capital losses of up to \$3,000 per year. Disallowed net capital losses can generally be used in other tax years to offset net capital gains. It also is common for net capital gains to face a lower tax rate than ordinary income. See, e.g., U.S. Internal Revenue Code section 1(h). This, however, may be viewed as relating to the fact that realization elasticity may tend to be higher for capital gains than ordinary income. For example, I may find it easier to continue holding appreciated stocks in my portfolio for tax reasons, than to adjust my labor supply.

increased to \$10 billion or \$100 billion, by simply adding zeroes to the swapped pieces of paper. Nonrefundability limits the potential economic benefit from engaging in such transactions, admittedly at an arbitrary point (i.e., when taxable income reaches zero), thus in effect capping the potential fiscal harm.

Disallowance for net capital losses responds to an even easier tax planning idea, which is to hold a well-diversified portfolio of individually risky financial assets, and then to sell the losers while holding the winners, thus creating a tax loss even if there is no overall economic loss. The incentive to engage in such strategic trading is, of course, inevitable in a realization-based system. Once again, denying net capital losses limits the net benefit available from engaging in this strategy, admittedly at an arbitrarily determined point.

Rules of this kind raise complicated efficiency issues. For example, they increase tax planning costs in some instances, by inducing taxpayers not to bother generating loss deductions, but increase such costs in others, by motivating efforts to circumvent the rules. Such issues lie at a considerable remove from the world of clean, analytically tractable economic models such as those from the OIT and optimal commodity tax literatures, but they cannot be ignored if one wishes to understand the main battlegrounds in a real world income tax system.

4 INCOME TAXATION AND CONSUMPTION TAXATION

For forty years, one theoretical issue has had pride of place in the U.S. academic tax policy literature. This is the question of whether the main tax instrument used in distribution policy should be an income tax or a consumption tax.¹⁸ The great majority of countries use both income taxes and value-added taxes (VATs) that generally treat current-year market consumption as the tax base. In the United States, however, with general consumption taxes

¹⁸ The U.S. academic debate concerning income and consumption taxation was effectively launched by Andrews (1974).

being limited to relatively small levies (retail sales taxes or RSTs) that are imposed at the state or local government level, much attention has focused on the idea of replacing the existing income tax with a progressive consumption tax. Such a tax would be partly or wholly collected from individuals or households, rather than from businesses like a VAT or RST, thus permitting the application of a graduated rate structure based on overall personal circumstances.¹⁹

Perhaps the main virtue of this debate, despite its narrow focus and apparently dim prospect of ever leading to the enactment of major policy changes, is that it has served as a powerful instrument for advancing theoretical understanding of intertemporal issues. The big difference between an income tax and a consumption tax is that the former reaches, while the latter excludes, normal returns to saving (which may be thought of as funding future consumption). By examining the significance of this difference, the income versus consumption tax debate has helped to focus attention on the significance of the time dimension in evaluating both efficiency and distribution, as well as on a variety of administrative and other tax base design issues.

In briefly reviewing the main contours of this vast, intricate, and multifaceted debate, it is useful to start with the famous Haig-Simons definition of economic income (set forth, for example, by Simons 1938, 50). If Y denotes income, C consumption, and ΔW change in net worth (i.e., the return to saving), then one's economic income for a given accounting period, such as a year, can be expressed as follows:

$$Y = C + \Delta W$$

¹⁹ In practice, existing VATs and RSTs typically leave a lot of consumption out of the tax base. In addition, in practice RSTs often reach business-to-business transactions, which do not involve consumption as properly defined.

A consumption tax, in contrast to an income tax, only reaches C . For example, a VAT or RST applies to consumer purchases.²⁰ However, one can also directly tax consumption at the household or individual level, without requiring people to keep track of all their consumer purchases, by subtracting ΔW from each side, and then rearranging the terms in the Haig-Simons equation as follows:

$$C = Y - \Delta W$$

In practice, this means that an individual or household-level measure of income that is modified to allow deductions for all saving, while also requiring inclusion of all dissaving, can reach the same consumption base as a VAT or RST. And with measurement taking place at the individual or household level, graduated rates can be imposed that reflect the taxpayer's overall circumstances (such as her overall consumption for the year, or over a longer period).²¹

What are the consequences of excluding ΔW from the tax base? If implemented naively, such as by excluding capital gains from the tax base even if they actually reflect labor income,²² people such as Bill Gates and Mark Zuckerberg would be able largely to exempt the huge fortunes that they have reaped via stock appreciation. However, the relevant tax policy literature shows that – at least with proper implementation and complete capital markets, and if one uses an unlimited time horizon to make the assessment – the difference between an income tax and a

²⁰ In a typical VAT, business-to-business sales are taxable to the seller, but the buyer gets an offsetting credit or deduction. The net tax imposed on business-to-business sales is therefore zero, just as under an RST, where they are disregarded.

²¹ Consumption taxes that use this methodology are known as cash-flow or consumed income taxes. The other main vehicle, within a consumption framework, for using individual or household-level information to guide the imposition of graduated marginal rates, is the flat tax or X-tax. This, in turn, is basically a flat-rate business-level VAT, modified to provide a business-level deduction for wages. The wages are then taxed to workers under a graduated rate structure. See Hall and Rabushka 1995; Bradford 2003; Carroll and Viard 2012.

²² Suppose that I found a new company, make it extremely valuable through my efforts, and then sell it for an enormous capital gain. From an economic standpoint, my profit may be viewed as reflecting labor income that should not be excluded from either an income tax or a consumption tax base.

consumption tax is limited to the *normal risk-free return* to saving, which the former, but not the latter, subjects to tax.

What do we mean by the “normal risk-free return?” This is what remains, out of the economic returns that people derive from or that are associated with their investment outlays, once one excludes the following three items, which the literature shows that an income tax and a consumption tax should in principle treat the same: (1) risk premia, (2) compensation for expected inflation, and (3) extraordinary returns, such as Bill Gates’ big score from founding Microsoft. One might think of the remaining residual as akin to the interest rate that a bank would offer on a secure savings account (allowing demand withdrawals, to eliminate interest rate risk) in the absence of expected inflation, and if the bank was not bundling its provision of financial services into the terms of deal.

While the full analysis that is needed to reach (or contest) this conclusion is beyond my scope here,²³ a simple example helps to make the distinct but related point that, with limited time and just a single rate of return (well below that which a Gates or Zuckerberg historically received on their main business ventures), a consumption tax, by reason of its excluding that return, is neutral as between current and future consumption. An income tax, by contrast, penalizes future consumption relative to current consumption.

To this end, suppose that the normal, risk-free rate of return is 10 percent. A taxpayer earns \$100 in Year 1, and can consume what she retains after-tax either immediately or in a year. Suppose that the tax rate, under either an income tax or a consumption tax, is 40 percent.

²³ For an overview of the logic that supports this analysis, see Shaviro 2004b, 100-103. Among the important contributions to the extensive literature that reached this conclusion are Domar and Musgrave 1944; Bankman and Griffith 1992; Kaplow 1994; and Bradford 1996.

Finally, suppose that the income tax is levied when the income is earned, while the consumption tax is levied at the time of consumption.²⁴

Consider first the income tax. After paying \$40 of tax on her earnings, the taxpayer can consume the remaining \$60 right away, in which case she will face no further income tax liability. Or she can save it for a year, earning \$6 of interest on which she pays a further \$2.40 of income tax. A decision to save would therefore increase her total income tax liability, not just nominally but in present value terms as measured up front. With a 10 percent interest rate, \$2.40 of income tax that is paid in a year has an up-front present value of \$2.18. Accordingly, adding this to the \$40 of tax that the taxpayer incurs up-front on her \$100 of earnings, one could view the income tax as imposing a 40 percent tax rate on immediate consumption, and a 42.18 percent tax rate on consumption that is deferred for a year.

Now consider the consumption tax. Just as under the income tax, consuming right away would generate \$40 of tax liability, permitting \$60 to be consumed. Suppose, however, that one instead waits for a year. At a 10 percent interest rate, one now has \$110, permitting the consumption of \$66 after paying \$44 of tax. Given the 10 percent interest rate, this amount has the same present value as \$40 paid immediately. Thus, current and future consumption face the same effective tax rate of 40 percent.

This point can be illustrated algebraically as follows. Suppose that X is the amount one earns, t is the tax rate, and r is the interest rate. Under a consumption tax, if one consumes immediately, then in the next period one has $X(1 - t)(1 + r)$. Consuming at the end of the next period leaves one with $X(1 + r)((1 - t))$. These terms are algebraically equivalent.

²⁴ At least in principle (whether or not feasibly in practice), it should be possible to change this timing method under either system without altering the present value of tax liabilities, and thus the incentive effects, of either system. See Shaviro 2000, 54-55.

Adding more years of deferral to the analysis does nothing to change the bottom line result. The effective income tax rate on deferred consumption continues to increase, as one adds further tax payments on continued saving. By contrast, the effective consumption tax rate remains the same no matter how long the taxpayer waits, since the amount that will ultimately be due grows at the discount rate.

To regard this as a pertinent description of how a consumption tax would actually bear on, say, Bill Gates or Mark Zuckerberg, one has to be willing to apply this reasoning to indefinite deferral. We don't know when (or perhaps even if) extremely rich people and their heirs will actually consume their huge fortunes. Yet the deferred tax burden does indeed keep growing at a market interest rate. Moreover, the deferred tax burden, even as it remains unpaid, does indeed affect the purchasing power of the wealth (which is, after all, what defines its value).

Accordingly, under a standard view of economic incidence, the consumption tax – so long as it will permanently remain in place – is indeed currently borne by wealth-holders, even if actual tax payments remain indefinitely deferred. However, not everyone finds this reasoning entirely satisfactory – whether because they feel a need for a finite “due date,” or based on the option value that savers may have to wait for a tax rate reduction. Accordingly, the fact that consumption tax payments, under a standard method of implementation, may remain indefinitely deferred stands, for many people, as a reason for preferring income taxation.

With or without such concerns, however, the fact that an income tax, unlike a consumption tax, discourages saving relative to immediate consumption has important efficiency implications under a tax neutrality framework. Given that an income tax effectively imposes higher tax rates on later-consumed than earlier-consumed items, the principle that, in general,

one should tax all consumer goods uniformly provides direct support for preferring consumption taxation to income taxation (Bankman and Weisbach 2006).

Other considerations could complicate the analysis, however. For example, suppose that an income tax is likely to be more progressive in practice, even though, in principle, a consumption tax could match it in progressivity if given more steeply graduated rates.²⁵ One who preferred greater progressivity might count this as an argument in favor of the income tax. In addition, purely from an efficiency standpoint, suppose that savings behavior is a signal or tag, associated high ability that cannot otherwise be observed (Saez 2002, 228). There also is literature suggesting that a positive tax on savings may be desirable because, “with high savings, one can more easily afford to under-utilize one’s earning ability by working less [thereby] impos[ing] a negative revenue externality” (Shaviro 2007, discussing the “new dynamic public finance literature,” such as Golosov et al 2003).

A final set of issues raised by the choice between income taxation and consumption taxation is administrative. In the words of William Andrews (1983, 280), a realization requirement may be the “Achilles’ heel” of any practically feasible income tax, including for the reasons that were briefly discussed in section 3.4 above. Consumption taxes can avoid the problems associated with realization, albeit facing various administrative, compliance, and enforcement challenges of their own.²⁶

5 ENTITY-LEVEL TAXATION OF CORPORATIONS

Among the most consequential implications of having a realization-based income tax pertains to income earned by corporations. Obviously, in the case of a pure Haig-Simons income

²⁵ Since high-earners tend to save a larger percentage of their earnings than low-earners, shifting from an income tax to a consumption tax with similar rate graduation would reduce the system’s progressivity if all else (including the degree of rate graduation) remained the same.

²⁶ On the administrative, compliance, and enforcement issues posed by the design of a progressive consumption tax, see, e.g., Weisbach 2009.

tax, all income would be taxed directly to individuals, and it would make no difference whether a corporation was used to conduct particular business operations. However, once we have a realization-based income tax system, the following considerations may support treating corporations as distinct taxpayers:

1) Corporations often have complex business operations, with management and record-keeping being centralized at the entity level, and with public reporting requirements that result in the preparation of company-level measures of financial accounting income. Under these circumstances, entity-level income taxation has large administrative advantages.

2) In principle, one could require corporations merely to report their income, with the entire amount being “flowed through” for tax purposes to individuals such as shareholders. (This is indeed how partnerships generally are taxed, in the U.S. and elsewhere.) However, several factors make this difficult to accomplish in a satisfying way. First, for companies with complex capital structures, it may be hard to divide the income between individuals in a reasonably accurate and not overly manipulable fashion. Second, financial interests that people hold outside of the companies’ internal capital structures – for examples, options and other derivative financial positions – may add to these difficulties. Third, to the extent that a corporation’s taxable income differs from its economic income, the question of *whose* income it is lacks any well-defined economic answer (Shaviro 2014, 181). Fourth, there may be political objections to requiring individuals to pay current taxes with respect to entity-level income that they cannot directly access absent control of the company.

3) While one could respond to these problems by imposing no tax on corporate income until it is directly realized by shareholders (such as through the receipt of dividends), this would

allow corporate investment to enjoy much broader income tax deferral than that which is generally available to prospective shareholders.

Taxing corporate income at the entity level almost inevitably has anomalous tax planning effects, such as encouraging or discouraging the use of corporate entities depending on whether the use of a corporation would lower or raise the applicable marginal tax rate. The resulting efficiency problems are exacerbated by longstanding features of “classical” corporate tax systems in the United States and elsewhere. In particular, such systems distinguish between financial instruments that are classified as corporate debt and equity, respectively. Corporate interest payments or accruals on instruments classified as debt are deductible, but dividends paid on equity are not.²⁷ This can both lead to double taxation of equity-financed corporate income and to discouraging dividend payouts.²⁸

Entity-level corporate income taxation also has important effects on the taxation of cross-border investment and capital flows. If all income earned through the use of corporations were taxed directly to the ultimate individual owners, it is plausible that countries would respond by taxing all outbound investment. After all, if income is being used as a proxy for ability or ability to pay, then presumably one should not care where a given individual’s earnings arose. It also is plausible that, absent the need for entity-level collection of the corporate income tax, inbound capital investment frequently would be exempted from source country taxation, given that inbound investors generally would not bear the incidence of a source-based tax if they could respond by investing elsewhere and getting the same after-tax return (see Shaviro 2009, 67-68).

²⁷ The U.S. tax rate imposed on the receipt of a dividend generally is 20 percent or lower, whereas the top rate for individuals’ ordinary income is 39.6 percent.

²⁸ Double taxation does not occur if the corporation can avoid entity-level taxes. In addition, if dividend taxation at a fixed rate is inevitable at some point, then, as the “new view” of dividend taxation makes clear, making payouts sooner rather than later is not discouraged by the tax. See Shaviro 2009, 73-77.

However, with corporate taxpayers' residence being determined at the entity level but in the face of cross-border shareholding, the available policy choices grow considerably murkier. Resident individuals' foreign source income cannot necessarily be taxed currently merely by imposing worldwide taxation on resident entities, since such individuals can invest abroad through foreign entities. In addition, source-based taxation of foreign entities is made costlier to forgo by the concern that resident individuals could hide behind such entities when earning income at home. All this contributes to a state of affairs in which (1) source-based corporate income taxation is nearly universal (other than in small countries that specialize in serving as tax havens), (2) the proper treatment of outbound investment by resident corporations is hotly contested, and (3) large multinational companies have opportunities to engage in large-scale tax avoidance that imposes high distortionary and tax planning costs relative to the taxes that they end up paying (see generally Shaviro 2014).

6 TAX REFORM

Few topics in Washington politics are more perennially popular in the abstract than fundamental or large-scale tax reform. While in principle "tax reform" could simply mean whatever changes to the tax system one happens to like, for several decades it has had a relatively fixed meaning in tax policy discourse. It has tended to denote something along the lines of the Tax Reform Act of 1986, in which Congress enacted an approximately budget-neutral (within the officially estimated period) and distributionally neutral package that combined base-broadening with rate reductions.

In recent years, the idea of attempting an at least modified 1986 replay has periodically attracted significant political attention. The big differences from 1986 are that some proposals would either (a) follow the 1986 script just for corporations, in response to other countries'

having lowered their corporate rates, or else (b) include converting the income tax into a consumption tax, or alternatively scaling it back and adding a VAT to replace the lost revenue.

Enthusiasm for 1986-style tax reform often wanes once people learn more about the tax benefits that would have to be repealed in order to finance significant rate cuts. In the U.S., for example, base-broadening for individuals might have to involve the repeal or curtailment of such popular items as home mortgage interest deductions, charitable deductions, and the exclusion for employer-provided health insurance. Base-broadening to pay for corporate rate cuts would place favorable tax rules for new investment on the chopping block, as well as raising awkward issues concerning the relative taxation of corporate and non-corporate businesses.

In part, the enduring conceptual appeal of 1986-style tax reform reflects aspirations concerning the achievability of political compromise. For example, distributional neutrality as between the pre- and post-reform tax code can facilitate legislative cooperation between those who disagree about distributional issues but have other areas of agreement about tax policy. Likewise, the enactment of a budget-neutral reform package at least ostensibly permits one to avoid enmeshing the tax reform debate with that concerning deficit reduction. Yet an overlap between the two debates is hard to avoid if base-broadening is being used to fund lower rates, rather than for some other purpose, such as reducing the need for future changes on the spending side of the federal budget.

The main efficiency advantage of 1986-style tax reform arises from base-broadening, in cases where this promotes neutrality with respect to rival consumption or investment choices that previously were taxed disparately. For example, repealing home mortgage interest deductions may reduce the income tax system's encouragement of home ownership relative to both rental occupancy and non-home consumption. However, the effect of cutting marginal rates in the

context of a 1986-style reform is more complicated. Rate reduction that was adopted as a stand-alone (and revenue-losing) change generally would be expected to reduce the deadweight loss resulting from the operation of the tax system. However, if this change is accompanied by base-broadening, such that overall income tax burdens remain about the same as previously, then the tax system's overall discouragement of work and saving should also remain about the same.²⁹

While 1986-style tax reform treats base-broadening and rate reduction as complements (based on assuming a fixed budgetary goal), there is a sense in which base-broadening and rate *increases* are the true complements. In general, the less avoidable a given tax instrument is, the less deadweight loss one would expect it to generate per dollar of revenue raised. At the limit, a lump-sum tax, by reason of being entirely unavoidable, generates zero deadweight loss.

The empirical literature concerning responses to income taxation suggests that labor supply tends to be less responsive to the tax rate, at least in the short run, than various other margins that reflect tax planning effort. Thus, base-broadening has the potential to reduce the efficiency cost of higher rates, despite its potentially adverse impact on labor supply. For example, the 70 percent top rate that Diamond and Saez (2011) advocate is more likely to be at or below the revenue-maximizing rate level if the tax base is otherwise broad than if it is narrow.

7 CONCLUSION

Within law and economics, tax law is relatively unusual in having distributional aims play a central role, alongside efficiency. Indeed, there is prominent literature asserting, although also contesting, the proposition that distributional issues should be left entirely to the tax and

²⁹ Suppose, for example, that a revenue-neutral reform cuts rates and repeals the home mortgage interest deduction, on balance increasing homeowners' tax burdens while reducing those of non-homeowners. The latter group has stronger work incentives than previously, but the former group may have weaker such incentives if, at the margin, they were earning money to finance home as well as non-home consumption.

transfer system, instead of having any influence on policy in other areas of law.³⁰ One consequence of this greater scope is that, at least within a conventional economics framework, tax policy issues may generally be both more complex and harder to resolve satisfyingly than those in many other areas of law. After all, tax requires both trading off efficiency concerns against distributional ones and evaluating efficiency in a very complex second-best framework.

To those who like to analyze issues that are relatively straightforward and have clear-cut answers, this greater complexity and indeterminacy may make tax policy an area to avoid. To others, they are part of its core appeal.

³⁰ This literature includes, for example, Kaplow and Shavell (1994) (asserting that the legal system is less efficient than the tax system at redistributing income) and Sanchirico (2000) (contesting this proposition).

REFERENCES

Andrews, W. 1974. A Consumption-Type or Cash-Flow Personal Income Tax. *Harvard Law Review* 87: 1113-1188.

Andrews, W. 1983. The Achilles' Heel of the Comprehensive Income Tax. In Walker, C. and Bloomfield, M. (eds.), *New Directions in Federal Tax Policy for the 1980s*. Pensacola, FL: Ballinger Publishing Co.

Atkinson, A.B. and Stiglitz, J.E. 1976. The Design of Tax Structure: Direct Versus Indirect Taxation. *Journal of Public Economics* 6: 55-75.

Auerbach, A. 2012. The Mirrlees Review: A U.S. Perspective. *National Tax Journal* 65: 685-708.

Bankman, J. and Griffith, T. 1992. Is the Debate Between an Income Tax and a Consumption Tax a Debate About Risk? Does It Matter? *Tax Law Review* 47: 377-___.

Bankman, J. and Weisbach, D. 2006. The Superiority of an Ideal Consumption Tax Over an Ideal Income Tax. *Stanford Law Review* 58: 1413-1456.

Bradford, D. 1996. Consumption Taxes: Some Fundamental Transition Issues. In Boskin, M. (ed.), *Frontiers of Tax Reform*. Stanford, CA: Hoover Institution Press.

Bradford, D. 2003. *The X-Tax in the World Economy*. Princeton, NJ: Princeton University Press.

Carroll, R. and Viard, A. 2012. *Progressive Consumption Taxation: The X Tax Revisited*. Washington: AEI Press.

Corlett, W.J., and Hague, D.C. 1953. Complementarity and the Excess Burden of Taxation. *Review of Economic Studies* 21: 21-30.

Diamond, P. and Saez, E. 2011. The Case for a Progressive Tax: From Basic Research to Policy Recommendations. *Journal of Economic Perspectives* 25: 165-190.

Domar, E. and Musgrave, R. 1944. Proportional Income Taxation and Risk-Taking. *Quarterly Journal of Economics* 58: 388-___.

Edgeworth, F. Y. 1897. The Pure Theory of Taxation: III. *Economic Journal* 7: 550-571.

Golosov, M., Kocherlakota, N., and Tsyvinski, A. Optimal Indirect and Direct Capital Taxation. *Review of Economic Studies* 70: 569-___.

Gruber, J. 2011. *Public Finance and Public Policy* (3rd ed.). New York: Worth Publishers.

Hall, R. and Rabushka, A. 1995. *The Flat Tax*. Stanford, CA: Hoover Institution Press.

Kaplow, L. 1994. Taxation and Risk Taking: A General Equilibrium Perspective. *National Tax Journal* 47: 789-___.

Kaplow, L. 2008. *The Theory of Taxation and Public Economics*. Princeton, NJ: Princeton University Press.

Kaplow, L. and Shavell, S. 1994. Why the Legal System is Less Efficient Than the Income Tax in Redistributing Income. *Journal of Legal Studies* 23: 667-___.

Kaplow, L. and Shavell, S. 2002. *Fairness Versus Welfare*. Cambridge, MA: Harvard University Press.

Mankiw, N.G., and Weinzierl, M. 2010. The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution. *American Economic Journal: Economic Policy*, 2(1): 155-76.

Mirrlees, J. 1971. An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* 38: 175-208.

Murphy, L. and Nagel, T. 2002. *The Myth of Ownership: Taxes and Justice*. Oxford: Oxford University Press.

Ramsey, F. P. 1927. A Contribution to the Theory of Taxation. *Economic Journal* 37: 47-61.

Saez, E. 2002. The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes. *Journal of Public Economics* 83: 217-230.

Sanchirico, C. 2000. Taxes Versus Legal Rules as Instruments for Equity: A More Equitable View. *Journal of Legal Studies* 29: 797-___.

Shaviro, D. 2000. *When Rules Change: An Economic and Political Analysis of Transition Relief and Retroactivity*. Chicago: University of Chicago Press.

Shaviro, D. 2004a. Rethinking Tax Expenditures and Fiscal Language. *Tax Law Review* 57: 187-___.

Shaviro, D. 2004b. Replacing the Income Tax with a Progressive Consumption Tax. *Tax Notes* 103: 91-113.

Shaviro, D. 2007. Beyond the Pro-Consumption Tax Consensus. *Stanford Law Review* 60: 745-788.

Shaviro, D. 2009. *Decoding the U.S. Corporate Tax*. Washington, D.C.: Urban Institute Press.

Shaviro, D. 2011. 1986-Style Tax Reform: A Good Idea Whose Time Has Passed. *Tax Notes* 131: 817-842.

Shaviro, D. 2014. *Fixing U.S. International Taxation*. New York: Oxford University Press.

Simons, H. 1938. *Personal Income Taxation*. Chicago: University of Chicago Press.

Slemrod, J. 1990. Optimal Taxation and Optimal Tax Systems. *Journal of Economic Perspectives* 4: 157-178.

Slemrod, J. and Yitzhaki, S. 2002. Tax Avoidance, Evasion, and Administration. In Auerbach, A. and Feldstein, M, (eds), *Handbook of Public Economics*, Volume 3, pages 1423-1470.

Weisbach, D. 2009. Implementing Income and Consumption Taxes. In Auerbach, A. and Shaviro, D. (eds.), *Institutional Foundations of Public Finance*. Cambridge, MA: Harvard University Press.