Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance

Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss?
The Implications of Evasion and Avoidance

Raj Chetty*

UC-Berkeley and NBER

August 2008

Abstract

Since Feldstein (1999), the most widely used method of calculating the excess burden of income taxation is to estimate the effect of tax rates on reported taxable income. Feldstein's taxable income formula for deadweight loss implicitly assumes that the marginal social cost of evasion and avoidance equals the tax rate. This paper argues that this condition is likely to be violated in practice for two reasons. First, some of the costs of evasion and avoidance are transfers to other agents in the economy rather than real resource costs. Second, some individuals overestimate the costs of evasion and avoidance. I show that, in such situations, excess burden depends on a weighted average of the taxable income and total earned income elasticities, with the weight determined by the resource cost of sheltering income from taxation. This generalized formula implies that the efficiency cost of taxing high income individuals is not necessarily large despite evidence that their reported incomes are highly sensitive to marginal tax rates.

Keywords: excess burden, tax evasion, optimal taxation

*E-mail: chetty@econ.berkeley.edu. I have benefited from discussions with Alan Auerbach, Peter Diamond, Caroline Hoxby, John Friedman, David Gamage, Roger Gordon, Louis Kaplow, Adam Looney, Wojtek Kopczuk, Emmanuel Saez, Joel Slemrod, Shlomo Yitzhaki, and Philippe Wingender. Gregory Bruich provided outstanding research assistance. Funding from the Hoover Institution is gratefully acknowledged.

In an influential pair of papers, Martin S. Feldstein (1995, 1999) showed that the excess burden of income taxation can be calculated by estimating the effect of taxation on reported taxable income – the "taxable income elasticity." Feldstein's taxable income approach has since become the central focus of the literature on taxation and labor supply because of its elegance and practicality. The approach is elegant because one does not have to account for the various channels through which taxation might affect behavior (e.g. hours, effort, training) to measure efficiency costs. It is practical because tax records containing data on reported taxable income are widely available.

The empirical literature on the taxable income elasticity has generally found that elasticities are large (0.5 to 1.5) for individuals in the top percentile of the income distribution, and relatively small (0 to 0.3) for the rest of the income distribution (see e.g., Lawrence B. Lindsey 1987, Joel B. Slemrod 1998, Jonathan Gruber and Emmanuel Saez 2002, Saez 2004). This finding has led some to suggest that reducing top marginal tax rates would generate substantial efficiency gains.[1]

The taxable income reported by high income individuals is very sensitive to the tax rate partly because of tax avoidance and evasion (Slemrod 1992, 1995).[2] For example, individuals make charitable contributions to reduce their taxable income or use unmonitored offshore accounts to under-report income. Does the efficiency cost of taxation depend on whether the taxable income elasticity is driven by avoidance and evasion rather than changes in labor supply? Existing studies (e.g. Feldstein 1999, Slemrod and Shlomo Yitzhaki 2002, Saez 2004) suggest that the answer is no, as long as there are no changes in tax revenue from other tax bases. For example, Slemrod and Yitzhaki remark that "Feldstein's (1999) claim about the central importance of the elasticity of taxable income generalizes to avoidance and evasion." The intuition underlying this conclusion is straightforward: an optimizing agent equates the marginal cost of sheltering $1 of income from taxation with the net marginal cost of reducing earnings by $1, so the reason that reported taxable income falls does not matter for efficiency calculations.

This paper reevaluates the taxable income elasticity as a measure of deadweight loss in the presence of evasion and avoidance ("sheltering" behaviors).[3] Feldstein's formula implicitly requires that the marginal *social* cost of sheltering $1 of income equals the tax rate (the benefit of sheltering

[1] Academic examples include Gruber and Saez (2002) and Feldstein (2006). The Joint Economic Committee (2001) has argued in favor of lowering top rates based on the taxable income evidence. See Austan Goolsbee (1999) for a critique of the empirical literature on taxable income.

[2] Income shifting can also occur intertemporally. When tax changes are anticipated, individuals appear to retime income substantially (Goolsbee 2000). I abstract from such intertemporal effects, focusing on the question of how to measure efficiency costs using estimates of the long-run effect of taxes on behavior.

[3] The distinction between illegal evasion and legal avoidance is not critical for the analysis in this paper, so I use the term "sheltering" as a general description of all evasion and avoidance behaviors.

\$1). This condition is likely to be violated in practice for two reasons. First, and most importantly, some of the costs of sheltering are transfers to other agents in the economy rather than real resource costs. For instance, an individual may be deterred from tax evasion because of the expected cost of being fined by the government or other agents in the private sector. Individuals seeking to avoid taxation by making charitable contributions or setting up trusts for their descendants may not fully internalize the benefits associated with their contributions, effectively incurring a transfer cost for sheltering income. Second, empirical studies have found that individuals overestimate the costs of sheltering – e.g. overestimating the detection probability and fines for tax evasion (James Andreoni, Brian Erard, and Jonathan Feinstein 1998). Such optimization errors also create a difference between the true marginal social cost of sheltering and the tax rate.

The taxable income formula for deadweight loss does not hold when the marginal resource cost of sheltering differs from the tax rate. Indeed, if sheltering has no resource costs, it generates no efficiency loss at all because it simply leads to a reallocation of resources across agents. In this case, deadweight loss depends purely on the total *earned* income elasticity – the effect of taxes on "real" choices that affect total earnings.

In the general case where sheltering has a positive resource cost that is not necessarily equal to the tax rate, I derive a simple formula for marginal excess burden that depends on a weighted average of the reported taxable income and total earned income elasticities. The weight is proportional to the loss in social surplus from of an additional dollar of sheltering – the marginal resource cost. Intuitively, reductions in total earnings caused by taxes always generate excess burden because they distort aggregate output. The additional excess burden generated by sheltering behaviors – which lead to a difference between earned and taxable income – is proportional to the marginal resource cost of such behaviors.

The formula developed here applies irrespective of whether transfer costs or optimization failures create the wedge between the marginal resource cost and the tax rate. The formula is also unaffected by revenue offsets (transfers to the government) that occur through shifting of income across tax bases. In this sense, the analysis generalizes Feldstein (1999) by providing a robust method of calculating marginal excess burden when the perceived private cost of sheltering differs from its social cost.

The results in this paper have several precedents in the literature, notably in the work of Slemrod and co-authors. It is widely recognized that the calculation of excess burden is complicated by revenue offsets in the presence of multiple taxes (e.g. Slemrod 1998, Roger H. Gordon and Slemrod

2002, Slemrod and Yitzhaki 2002, Alan J. Auerbach and James R. Hines Jr. 2002, Saez 2004). Slemrod (1998) and Saez (2004) propose formulas that adjust for revenue offsets by adding terms for the change in revenue from other taxes. In addition, Slemrod (1995) and Slemrod and Yitzhaki (2002) observe that fines lead to a difference between the private and social costs of evasion. They note that this difference will create an added term that must be taken into account when calculating excess burden, but do not characterize that term formally.

This paper contributes to the taxable income elasticity literature in three ways. First, it shows how transfers between agents within the private sector affect the calculation of excess burden. Existing formulas that adjust for revenue offsets are not valid in the presence of private transfers. Second, unlike earlier formulas, the formula derived here accommodates optimization errors in sheltering. Third, even ignoring within-private-sector transfers and optimization errors, the formula offers an alternative approach to measuring the marginal excess burden of taxation in the presence of revenue offsets. This alternative representation permits calculation of the excess burden of an income tax change without characterizing its effects on other tax bases, as would be required to implement the Slemrod and Saez formulas. It also yields some new intuition into the key determinants of excess burden. For instance, in the extreme case of pure transfer costs, the analysis shows that sheltering has zero efficiency cost when the added term mentioned by Slemrod and Yitzhaki is taken into account, completely severing the link between the taxable income elasticity and excess burden.

The results in this paper point to the marginal resource costs of avoidance and evasion as key parameters to be estimated in empirical studies of taxation. Since the resource costs of sheltering could potentially be much smaller than top marginal tax rates, one cannot conclude that the efficiency cost of taxing high income individuals is large directly from existing evidence of large taxable income elasticities.

I Theoretical Analysis

This section presents formulas for the marginal excess burden of a linear income tax under various assumptions about the costs of sheltering. As a reference, I first derive Feldstein's (1999) formula in a model without sheltering. Section I.B considers a model where individuals can avoid or evade taxes by paying a real resource cost. In section I.C, I analyze a model where sheltering has no resource cost but requires a transfer to another agent. Section I.D presents a general formula for marginal excess burden when sheltering has both resource and transfer costs. In section I.E, I show that this general formula is unaffected by optimization errors in sheltering decisions. Finally, section I.F considers the implications of the efficiency analysis for optimal taxation. To simplify the exposition, I abstract from income effects by assuming quasilinear utility in the main text. In Appendix A, I analyze a model with curved utility and show that the same formula is obtained with the uncompensated elasticities replaced by compensated elasticities.

I.A Benchmark Model: No Sheltering

Consider the canonical static labor-leisure model, where an individual chooses how many hours to work (l) at a fixed wage rate w. Let t denote the tax rate on labor income, y unearned income, c consumption, and $\psi(l)$ the disutility of labor. Let $R(t) = twl$ denote tax revenue. The individual's problem is to

$$\begin{aligned} \max_l u(c,l) &= c - \psi(l) \\ \text{s.t. } c &= y + (1-t)wl \end{aligned} \tag{1}$$

As is standard in excess burden calculations, the conceptual experiment I consider is to measure the net dollar-value loss from raising the tax rate and returning the revenue lump-sum to the taxpayer. For this purpose, I define social welfare as the sum of the individual's utility (which is a money metric given quasilinearity) and tax revenue:

$$W(t) = \{y + (1-t)wl - \psi(l)\} + twl \tag{2}$$

Since the agent has chosen l to maximize utility, the envelope condition implies that an increase in t has only a mechanical first-order effect on the agent's utility (i.e. $\frac{du}{dt} = \frac{\partial u}{\partial t} = -wl$). Hence, behavioral responses can be ignored when differentiating the term in the curly brackets, yielding the following expression for the marginal excess burden of taxation:

$$\begin{aligned} \frac{dW}{dt} &= -wl + wl + t\frac{d[wl]}{dt} \\ &= t\frac{dTI}{dt} \end{aligned} \tag{3}$$

where $TI = wl$ is taxable income. Feldstein (1999) showed that $\frac{dW}{dt} = t\frac{dTI}{dt}$ even in a model where individuals make a vector of decisions (l_1,...,l_n), such as hours, training, and occupation. Thus, the taxable income elasticity ($\frac{dTI}{dt}$) is a sufficient statistic to calculate deadweight loss in a general multi-input labor supply model. I return to the multi-input case in greater detail in Section II.A.

I.B Sheltering with a Resource Cost

Suppose the individual can shelter e dollars of income from taxation by paying a resource cost $g(e)$. The sheltering of e could be accomplished either through legal tax avoidance (e.g. setting up a trust) or illegal tax evasion (e.g. under-reporting). The cost $g(e)$ reflects the economic opportunity cost of sheltering e, i.e. the loss in total output from this behavior. For example, $g(e)$ could reflect the loss in profits from transacting in cash instead of electronic payments or the cost of choosing a distorted consumption bundle to avoid taxes.

The individual now chooses both labor supply (l) and how much income to shelter (e):

$$\begin{aligned} \max_{l,e} u(c,l,e) &= c - \psi(l) - g(e) \\ \text{s.t. } c &= y + (1-t)(wl - e) + e \end{aligned} \tag{4}$$

Social welfare is:

$$W(t) = \{y + (1-t)(wl - e) + e - \psi(l) - g(e)\} + t(wl - e) \tag{5}$$

Since the individual chooses l and e to maximize the term in curly brackets, we can again ignore behavioral responses when differentiating the term in the curly brackets. Hence we again obtain the Feldstein formula:

$$\frac{dW}{dt} = -[wl - e] + [wl - e] + t\frac{d[wl - e]}{dt} = t\frac{dTI}{dt} \tag{6}$$

where $TI = wl - e$ is reported taxable income. From an efficiency perspective, it does not matter if taxable income falls with t because of a change in labor supply (l) or a reporting effect (e). Intuitively, the agent optimally sets the marginal cost of reporting \$1 less to the tax authority ($g'(e)$) equal to the marginal private value of doing so (t). Since the agent supplies labor up to the point where his marginal disutility of earning another dollar equals $1 - t$, the marginal social value of earning an extra dollar (net of disutility of labor) is t. Hence, the marginal social costs of reducing earnings and reporting less income are the same at the individual's optimal allocation, making it irrelevant for efficiency which mechanism underlies the change in TI. This is the intuition underlying studies which argue that the taxable income elasticity is sufficient to calculate deadweight loss even in the presence of evasion and avoidance.

I.C Sheltering with a Transfer Cost

Part of the cost of sheltering may reflect a transfer across agents. The leading example of such transfer costs are fines levied for tax evasion. A fine has a private cost to the tax evader but has no social cost if the agent is risk neutral, because it simply redistributes output from the agent to the government.[4] If the agent is risk averse, the disutility associated with the increased uncertainty created by random audits constitutes a resource cost. To simplify the exposition, I analyze the risk-neutral case in this subsection and also assume that audits have no administrative cost. In Appendix A, I show that the general formula for excess burden in section I.D accommodates these factors if the utility cost of added risk and administrative cost of audits are included when measuring the marginal resource cost $g'(e)$.

I model audit deterrence of evasion using a simple variant of the framework developed by Michael G. Allingham and Sandmo (1972). Suppose that an individual is audited with probability $p(e)$, where $dp/de > 0$ – egregious under-reporting may lead to a higher chance of being caught. If caught, the individual must pay his tax bill plus a fine $F(e,t)$. Let $z(e,t) = p(e)[te + F(e,t)]$ denote the expected private cost of evasion.[5] Assume that z is a strictly convex function of e and that $\frac{\partial z}{\partial e}(0,t) = 0$ and $\frac{\partial z}{\partial e}(wl,t) = \infty$ to guarantee an interior optimum in e. Aside from these regularity assumptions, the derivation that follows does not depend on the specification of $z(e,t)$. Hence, although I have given a micro-foundation for $z(e,t)$ in terms of auditing for concreteness, the results below apply for any transfer cost z, i.e. any cost of sheltering which has a positive externality of equal size on another agent. I give other examples of transfer costs in section II.B.

The individual chooses e and l to

$$\begin{aligned} \max_{e,l} u(c,l,e) &= c - \psi(l) \\ \text{s.t. } c &= y + (1-t)(wl-e) + e - z(e,t) \end{aligned} \tag{7}$$

This agent's problem is formally identical to that in (4). However, there is a key difference in the social welfare function, in which the $z(e,t)$ transfer externality now appears twice (with opposite

[4] Only cash fines are transfers; imprisonment generates a social cost both for the criminal and the government, which must maintain the prison. The formula derived in section II.D allows for such resource costs of punishment.

[5] If $z(e,t)$ is linear in t (i.e. $z(e,t) = tz(e,1)$ for all t), the tax rate has no effect on sheltering: $\frac{de}{dt} = 0$ (Yitzhaki 1974). The excess burden calculations accommodate this case as well as any other specification of $z(e,t)$ because the properties of the cost function z are captured in the empirically estimated elasticities.

signs):

$$W(t) = \{y + (1-t)(wl - e) + e - z(e,t) - \psi(l)\} + z(e,t) + t(wl - e) \tag{8}$$

Exploiting the now familiar envelope condition for the term in curly brackets yields

$$\begin{aligned} \frac{dW}{dt} &= -(wl-e) - \frac{\partial z}{\partial t} + (wl-e) + \frac{\partial z}{\partial t} + \frac{\partial z}{\partial e}\frac{de}{dt} + t\frac{d[wl-e]}{dt} \\ &= t\frac{dTI}{dt} + \frac{\partial z}{\partial e}\frac{de}{dt} = t\frac{dLI}{dt} + \frac{de}{dt}(\frac{\partial z}{\partial e} - t) \end{aligned} \tag{9}$$

where $LI = wl$ is total (pretax) earned income. To simplify this expression, consider the first-order condition for the individual's choice of e:

$$t = \frac{\partial z}{\partial e} \equiv z'(e) \tag{10}$$

Intuitively, the individual sets the marginal private benefit of raising e by \$1 (saving \$$t$) equal to the marginal private cost. Combining (10) with (9) yields

$$\frac{dW}{dt} = t\frac{dLI}{dt} \tag{11}$$

Equation (11) shows that, in the transfer cost model, the taxable income elasticity is not a sufficient statistic to calculate deadweight loss. Instead, excess burden is determined purely by the effect of taxation on *total* earned income ($\frac{dLI}{dt}$) – the effect of taxes on "real" labor supply behavior.[6] If TI responds to t only because of sheltering, there is no deadweight loss. Intuitively, the total size of the pie is unaffected by sheltering in this model – the transfer cost z simply affects how the pie is split. In contrast, in the resource cost model, the cost of sheltering $g(e)$ is pure waste.

I.D Resource and Transfer Costs

Now suppose that sheltering e dollars of income from taxation requires payment of both a resource cost $g(e)$ and a transfer cost $z(e,t)$. The individual chooses e and l to maximize his utility net of

[6] Slemrod (2001) proposes a model of evasion in which the expected fine z depends on earnings (wl) as well as the amount evaded (e). For example, earning a higher income could increase the probability of audit or reduce the cost of hiding income. The qualitative results in this paper hold when $\frac{\partial z}{\partial wl} \neq 0$. However, (11) has an added term $\frac{\partial z}{\partial wl}\frac{\partial wl}{\partial t}$, reflecting the effect of earned income on the transfer cost. This is because $\psi'(l) = 1 - t - \frac{\partial z}{\partial wl}\frac{\partial wl}{\partial t}$ at the optimum, changing the social cost of distortions in labor supply.

both resource and transfer costs:

$$\max_{e,l} u(c,l,e) = c - \psi(l) - g(e) \tag{12}$$
$$\text{s.t. } c = y + (1-t)(wl-e) + e - z(e,t)$$

Social welfare is

$$W(t) = \{y + (1-t)(wl-e) + e - z(e,t) - \psi(l) - g(e)\} + z(e,t) + t(wl-e) \tag{13}$$

The same derivation as in the previous subsection gives

$$\frac{dW}{dt} = t\frac{dLI}{dt} + \frac{de}{dt}(\frac{\partial z}{\partial e} - t) \tag{14}$$

The difference from the pure transfer cost model is that the first order condition now has an additional term relative to (10), reflecting the marginal resource cost:

$$t = z'(e) + g'(e) \tag{15}$$

This leads to a general formula for deadweight burden with transfer and resource costs:

$$\frac{dW}{dt} = t\frac{dLI}{dt} - g'(e)\frac{de}{dt} \tag{16}$$
$$= t\{\mu\frac{dTI}{dt} + (1-\mu)\frac{dLI}{dt}\} \tag{17}$$
$$= -\frac{t}{1-t}\{\mu TI\varepsilon_{TI} + (1-\mu)wl\varepsilon_{LI}\} \tag{18}$$

where $\mu = \frac{g'(e)}{t} = \frac{g'(e)}{z'(e)+g'(e)}$ denotes the fraction of the total cost of sheltering accounted for by resource costs and $\varepsilon_{TI} = -\frac{dTI}{dt}\frac{1-t}{wl-e}$ and $\varepsilon_{LI} = -\frac{dLI}{dt}\frac{1-t}{wl}$ denote the taxable income and total earned income elasticities, respectively. When there is no transfer cost, $\mu = 1$, and (18) reduces to the standard taxable income formula for deadweight loss in (6). At the other extreme, when there is no resource cost, $\mu = 0$, and the formula reduces to (11), which depends only on the earned income elasticity. In the general case where $\mu \in (0,1)$, the marginal excess burden is determined by a weighted average of the taxable income and total earned income elasticities. The weight μ is determined by the magnitude of resource costs of sheltering relative to the tax rate or, equivalently, the magnitude of resource costs relative to the total (resource plus transfer) costs of sheltering.

Estimating the earned income elasticity ε_{LI} requires a method of inferring total earnings empirically. Avoidance behaviors that generate tax deductions – such as contributions to a charity or trust – are generally reported on income tax returns. One can therefore construct a gross-of-avoidance measure of earned income from the tax data used to estimate the taxable income elasticity. Recovering total earnings in the presence of evasion requires additional data. Total earnings can be imputed from consumption data, as in Christopher A. Pissarides and Guglielmo Weber (1989). One can also directly estimate the effect of tax rates on evasion ($\frac{\partial e}{\partial t}$) using audit data, as in Charles T. Clotfelter (1983). If sheltering responses are much larger than "real" labor supply responses – as documented by Slemrod (1992, 1995) – an estimate of ε_{LI} is less important because excess burden can be approximated simply by multiplying ε_{TI} by μ.

Why is it necessary to distinguish evasion and avoidance responses from changes in labor supply to calculate excess burden when $\mu < 1$? The agent's choice of l equates the marginal social cost of reducing labor supply, $w - \psi'(l)$, with the tax rate t. However, the agent's choice of e equates the marginal *private* cost of sheltering $(g' + z')$ with the tax rate t. Because the marginal private cost of sheltering differs from its social cost when $\mu < 1$, the key condition underlying Feldstein's formula – that the marginal social cost of all behaviors equals the tax rate – is violated for sheltering. This forces us to separate the two behaviors to calculate excess burden. Further intuition for why the violation of the $g'(e) = t$ condition makes $\frac{dW}{dt}$ a weighted average of ε_{TI} and ε_{LI} is given in section II.A.

Abstractly, it is not surprising that transfer costs affect the formula for excess burden, since externalities always introduce additional terms in efficiency calculations. Transfers are a subset of externalities where the agent making the sheltering decision bears a cost that is fully offset by a positive externality on other agents. Sheltering can also generate non-transfer externalities, such as the costs borne by the government for audits or imprisonment of tax evaders. Such non-transfer externalities affect the social resource costs of sheltering. Equation (18) holds with non-transfer externalities as long as they are included when measuring $g'(e)$.[7] It does not matter whether the individual making the sheltering decision or other agents bear resource costs, because social welfare $W(t)$ depends only on total resource costs to society.

[7] Unlike transfer externalities, however, negative non-transfer externalities generated by sheltering would make the standard taxable income formula in (3) *understate* deadweight loss.

I.E Optimization Errors

The analysis thus far has highlighted transfer costs as a source of a wedge between marginal resource costs of sheltering and the tax rate. Such a wedge can also arise from misperceptions of the cost of sheltering. Surveys show that many individuals substantially overestimate audit rates and fines associated with tax evasion (John T. Scholz and Neil Pinney 1993; Dick J. Hessing et al. 1992; Andreoni, Erard, and Feinstein 1998). Misinformed agents may not take full advantage of sheltering strategies despite their low marginal resource costs.

Motivated by this evidence, I extend the analysis to allow for misperceptions and optimization errors in sheltering decisions. Suppose the agent perceives the resource cost of sheltering to be $\widehat{g}(e)$ and the transfer cost of sheltering to be $\widehat{z}(e,t)$. For example, in the auditing model of section II.C, if the agent's perceived probability of being caught for evasion is $\widehat{p}(e)$ and his perceived fine is $\widehat{F}(e)$, then $\widehat{z}(e,t) = \widehat{p}(e)[te + \widehat{F}(e)]$. The individual chooses l and e to maximize his perceived expected utility

$$\max_{e,l} u(c,l,e) = y + (1-t)(wl-e) + e - \widehat{z}(e,t) - \psi(l) - \widehat{g}(e) \tag{19}$$

Social welfare is unaffected by the individual's perceptions and remains the same as in (13):

$$W(t) = \{y + (1-t)(wl-e) + e - z(e,t) - \psi(l(t)) - g(e)\} + z(e,t) + t(wl-e) \tag{20}$$

Because the utility function is separable in l and e, l is effectively chosen to optimize the term in the curly brackets even though e is not. Using the envelope condition for l and recognizing that e is not optimized, we obtain

$$\begin{aligned} \frac{dW}{dt} &= -(1-t)\frac{de}{dt} + \frac{de}{dt} - g'(e)\frac{de}{dt} + t(\frac{dLI}{dt} - \frac{de}{dt}) && (21)\\ &= t\frac{dLI}{dt} - g'(e)\frac{de}{dt} && (22)\\ &= t\{\mu\frac{dTI}{dt} + (1-\mu)\frac{dLI}{dt}\} && (23) \end{aligned}$$

where $\mu = \frac{g'(e)}{t}$.[8] Equation (23), which is the most general formula for marginal excess burden presented in this paper, coincides with the formula obtained above in (18). The intuition is again

[8] When agents do not optimize, one cannot write $\mu = \frac{g'(e)}{g'(e)+z'(e)}$ because the first order condition $g'(e) + z'(e) = t$ need not hold. Hence, resource costs should be normalized by the tax rate to calculate the weight in the general case where agents face transfer costs and do not optimize perfectly.

that the agent does not equate the marginal social cost of sheltering with the tax rate, making it necessary to weight $\frac{de}{dt}$ by $g'(e)$ instead of t when calculating marginal excess burden.

Since (23) does not rely on any assumptions about $\widehat{g}(e)$ or $\widehat{z}(e,t)$, it holds irrespective of the agent's perceived transfer and resource costs. More generally, one does not have to specify the positive model of behavior that drives the choice of e in order to derive (23). Hence, provided that we can measure $g'(e)$, we do not need to have a complete explanation of why there is a gap between $g'(e)$ and t to calculate marginal excess burden. This is an especially useful property because the model that explains observed evasion and avoidance behavior is debated (Andreoni, Erard, and Feinstein 1998).

Several caveats must be kept in mind in implementing (23) when agents make optimization errors. First, the labor supply decision must be separable from the sheltering decision to obtain (23) – that is, the optimal choice of l must be invariant to the choice of e. Intuitively, l will not be set at the unconstrained optimum if e is not optimized and sheltering affects the marginal return to work.[9] Second, recent evidence suggests that individuals misperceive not only the fines for tax evasion but also tax rates themselves. If agents choose l suboptimally, the first order condition for l does not hold and hence the $t\frac{dLI}{dt}$ component of (23) requires modification. The formula can be extended to accommodate optimization errors with respect to tax rates using an estimate of the wage elasticity of labor supply ($\frac{dLI}{dw}$), as shown by Chetty, Adam Looney, and Kory Kroft (2007). Finally, although a fully specified model for why $g'(e)$ differs from t in equilibrium is not required, some understanding of the positive model that drives sheltering is needed to measure $g'(e)$. For example, if agents do not evade taxes because of private ethical costs of doing so, one may include these costs when calculating $g'(e)$. In contrast, if agents do not evade taxes because of misinformation about audit rates and fines, the actual marginal resource cost $g'(e)$ would be lower.

I.F Implications for Optimal Taxation

The preceding results have different implications for the measurement of deadweight loss and the determination of optimal taxes. The taxable income elasticity is always a necessary input for revenue and optimal tax calculations, irrespective of its relevance for excess burden calculations. To illustrate this point, I consider a simple Ramsey tax problem. Suppose that one dollar of

[9] For example, separability is violated if the cost of evasion z depends on earnings wl or if there are income effects in labor supply (Yitzhaki 1987).

government spending on public goods generates social benefits of $1 + \lambda$, so that social welfare is

$$\widetilde{W}(t, \lambda) = \{y + (1 - t)(wl - e) + e - z(e, t) - \psi(l(t)) - g(e)\} + z(e, t) + (1 + \lambda)(wl - e)t \tag{24}$$

This social welfare function nests that in sections I.A-I.E, where I assumed $\lambda = 0$ in order to reproduce the compensating-variation measure of excess burden. When $\lambda = 0$, the optimal tax rate is trivially $t^* = 0$ since taxation generates no benefit but creates an efficiency cost. The interesting case for optimal taxation is the situation where $\lambda > 0$. A derivation analogous to that above gives

$$\frac{d\widetilde{W}}{dt} = -\frac{t}{1-t}\{\mu TI\varepsilon_{TI} + (1 - \mu)wl\varepsilon_{LI}\} + \lambda TI(1 - \frac{t}{1-t}\varepsilon_{TI}) \tag{25}$$

It follows that the tax rate t^* that maximizes $\widetilde{W}(t)$ satisfies:

$$\frac{t^*}{1-t^*} = \frac{\lambda}{(\mu + \lambda)\varepsilon_{TI} + (1 - \mu)\frac{wl}{wl-e}\varepsilon_{LI}} \tag{26}$$

When $\mu = 1$, (26) collapses to the standard inverse-elasticity rule for optimal taxation ($\frac{t^*}{1-t^*} = \frac{\lambda}{1+\lambda}\frac{1}{\varepsilon_{TI}}$). When $\mu = 0$, t^* is a function of both ε_{TI} and ε_{LI}, even though excess burden depends purely on ε_{LI}. This is because the relevant consideration for optimal tax design is the marginal cost of public funds – the deadweight cost generated per dollar of revenue raised:

$$MCPF(t) = -\frac{dW/dt}{dR/dt} = \frac{-\frac{t}{1-t}\{\mu TI\varepsilon_{TI} + (1 - \mu)wl\varepsilon_{LI}\}}{(wl - e)(1 - \varepsilon_{TI}\frac{t}{1-t})} \tag{27}$$

The optimal linear tax t^* equates the marginal cost of public funds with the marginal benefit of public expenditure: $MCPF(t^*) = \lambda$.[10] The taxable income elasticity determines how much the tax rate must be raised to generate an additional dollar of tax revenue (the denominator of (27)), while the earned income elasticity affects the marginal efficiency cost of that tax increase (the numerator of (27)). The formula for $\frac{dR}{dt}$ is unaffected by transfer costs or errors in optimization, and depends on ε_{TI} irrespective of μ. Thus, estimates of both ε_{TI} and ε_{LI} are required to analyze optimal taxation, even if sheltering has zero resource cost.[11]

[10] The MCPF is given by (27) only in a Ramsey model with a linear tax. In non-linear income tax models, the formula for the MCPF is more complex, as shown by Dahlby (1998). However, the qualitative point that both ε_{TI} and ε_{LI} matter for optimal taxation applies in models of non-linear taxation.

[11] A related point is that compliance costs borne by the individual must be distinguished from administrative costs of tax collection in determining t^*, because administrative costs enter the denominator of the $MCPF$ whereas

Note that even if sheltering has no efficiency cost, it could still be desirable to reduce sheltering from the perspective of optimal policy. Reducing tax evasion through stiffer penalties could be a more efficient way to generate revenue than raising distortionary taxes even if evasion has no resource cost (Louis Kaplow 1991, Joram Mayshar 1991). Again, additional factors beyond the marginal excess burden of raising t are relevant for the optimal policy problem.

II Discussion

II.A Foundations of Weighted Average Formula

Why does relaxing the assumption that $g'(e) = t$ lead to a formula for marginal excess burden that is a weighted average of ε_{TI} and ε_{LI}? To obtain an answer this question, it is useful to analyze the efficiency cost of income taxation at a more abstract level. Consider a model where the agent makes N choices $\{x_1, ..., x_N\}$ that contribute additively to taxable income, so that total taxable income is $TI = \sum x_i$. Choice x_i has a convex, increasing social cost of $g_i(x_i)$. The government levies an income tax t on taxable income, so that social welfare is given by

$$W = \{(1-t)\sum x_i - \sum g_i(x_i)\} + t\sum x_i \tag{28}$$

In this model, the excess burden of increasing the income tax t can be expressed as the sum of each of the behavioral responses weighted by their marginal social costs:

$$\frac{dW}{dt} = \sum_{i=1}^{N} g_i'(x_i)\frac{dx_i}{dt} \tag{29}$$

This expression for marginal excess burden is the most robust available because it does not rely on agent optimization and permits arbitrary externalities in the N decisions. The shortcoming of this formula is that it is difficult to separately estimate all behavioral responses and their total marginal social costs. Empirical implementation can be simplified by making assumptions about the underlying positive model. Feldstein (1999) assumes that there are no externalities and that agents choose $\{x_i\}_{i=1}^N$ optimally, so that $g_x'(x_i) = t\ \forall i$. Under this assumption, $\frac{dW}{dt} = t\sum_{i=1}^N \frac{dx_i}{dt} = t\frac{dTI}{dt}$.

The present paper makes a weaker assumption: choices that affect total earnings (LI) are made

compliance costs do not (Slemrod and Yitzhaki 1996).

optimally and do not generate externalities, but choices that create a difference between total earnings and reported taxable income ($TI - LI$) may be suboptimal and may generate externalities. Suppose that choices 1 to n affect total earnings, so that $LI = \sum_{i=1}^{n} x_i$ and choices $n+1$ to N affect only reported taxable income, so that total sheltering is $e = TI - LI = -\sum_{i=n+1}^{N} x_i$. Then

$$\frac{dW}{dt} = \sum_{i=1}^{n} g_i'(x_i)\frac{dx_i}{dt} + \sum_{i=n+1}^{N} g_i'(x_i)\frac{dx_i}{dt} \tag{30}$$

$$= t\sum_{i=1}^{n} \frac{dx_i}{dt} + \sum_{i=n+1}^{N} g_i'(x_i)\frac{dx_i}{dt} \tag{31}$$

$$= t\frac{dLI}{dt} - g'(e)\frac{de}{dt} \tag{32}$$

where $g'(e) = \frac{\sum_{i=n+1}^{N} g_i'(x_i)\frac{dx_i}{dt}}{\sum_{i=n+1}^{N} \frac{dx_i}{dt}}$ denotes the mean marginal resource cost of sheltering.

This derivation shows that the weaker assumption made here about the decisions underlying LI and TI directly leads to the weighted average formula for $\frac{dW}{dt}$ in (23). It also shows that (23) remains valid when earnings are affected by multiple decisions (e.g. training, effort, occupation) because ε_{LI} automatically aggregates all these behavioral responses. If these labor supply choices also have externalities, they should also be treated like sheltering decisions. For example, parents' choice of labor supply may have an externality on children or executives' leisure choices may affect their assistants' leisure. In these cases, one should weight $\frac{dLI}{dt}$ by its average marginal social cost including all external effects, instead of the tax rate, to calculate $\frac{dW}{dt}$. Thus, the formula for marginal excess burden proposed here is accurate when the important differences between social and perceived private costs are in sheltering decisions rather than total earnings behavior.

II.B Examples of Transfer Costs

There are two types of transfer costs: transfers to the government (revenue offsets) or to other agents in the private sector. Fines for tax evasion are the simplest example of transfers to the government. A more important example in practice is the shifting of income between tax bases, e.g. from personal to corporate income (Gordon and Slemrod 2002). If corporate income is taxed at a rate t_c, the agent effectively pays a transfer cost of $z(e) = t_c e$ to the government to shelter e from income taxation through shifting.

Private transfer costs can arise from penalties for evasion and avoidance imposed by other agents in the private sector. For example, a manager may be fired by shareholders if he is discovered to be using illegal tax sheltering strategies, thereby losing a bonus. A firm may lose clients to a competitor if it is identified as a tax evader. An individual may lose his wealth to theft by holding it in the form of cash or hidden accounts. Penalties that deter evasion could also deter quasi-legal avoidance strategies – e.g. declaring a vehicle or travel expense as a "business expense" – since the border between avoidance and evasion is often ambiguous. Misperceptions of such penalties would increase perceived transfer costs.

Avoidance strategies that are clearly legal can also be deterred by private transfer costs. The most important example is charitable contributions, whose benefits to recipients may not be fully internalized by the donor, effectively creating a transfer cost (Kaplow 1995). Empirical studies have shown that charitable contributions are highly sensitive to tax rates (see e.g. Feldstein and Amy Taylor 1976, Clotfelter 1985), suggesting that a significant part of the taxable income response observed in the data could be driven by this margin. Transfers within a family are another example. If an individual values his children's consumption at less than $1 but the planner weights all individuals equally, the individual effectively incurs a transfer cost when sheltering money from taxation through trusts or inter-vivos transfers.[12]

Transfer costs can also arise indirectly within the private sector. Suppose an executive is deciding whether to compensate himself in the form of a taxed dollar of labor income or untaxed perks such as amenities in the office (e.g. a better building, child care facilities, better company cars). Such perks have two forms of transfer costs. First, since an executive typically cannot take the perks with him to another job, some fraction of the benefits (in expectation) are transferred to subsequent executives. Second, some of the surplus from the office amenities may have to be shared with other employees even contemporaneously – it is difficult to improve only part of a building.

In considering the effects of indirect transfers, it is important to note that the only private transfer externalities that affect the Feldstein (1999) formula are those which are not internalized by agents through Coasian bargaining. For example, if a manager renegotiates his contract with those affected by his sheltering behavior (shareholders, employees, etc.), he effectively faces no private transfer cost in sheltering because his salary can be adjusted to offset any externalities. In practice, transaction costs and information problems likely impede perfect state-contingent contracting, and some sheltering behaviors may therefore be deterred by indirect private transfer costs at the margin.

[12] Each of these examples also involves some resource costs. For instance, the efforts spent by charities or children on lobbying to receive a transfer constitutes a resource cost.

Moreover, Coasian bargaining cannot overcome the externalities involved in direct redistribution of money to charities or family (Kaplow 1995).

II.C Comparison to Existing Formulas

The formula proposed here is not the only method of accounting for government transfer costs (revenue offsets) when agents optimize fully. An alternative approach is to compare the mechanical change in total tax revenue (from all tax bases) with the actual change in total tax revenue (Auerbach 1985, Slemrod 1998):

$$\frac{dW}{dt} = \frac{dR}{dt} - \frac{\partial R}{\partial t}|_{l,e} \tag{33}$$

where $R = t(wl - e) + z(e)$ denotes total tax revenue from all tax bases (including fines collected from audits) and $\frac{\partial R}{\partial t}|_{l,e}$ denotes the mechanical change in tax revenue if l and e were unchanged.

Saez (2004) proposes a variant of the revenue-distortion approach for the special case where agents shelter income by shifting earnings from the income to corporate tax bases. Saez's approach is to adjust Feldstein's (1999) formula by adding a term reflecting the added revenue raised from the corporate tax via shifting:

$$\frac{dW}{dt} = -\{\frac{t}{1-t}(wl-e)\varepsilon_{TI} - \frac{t_c}{1-t}e\varepsilon_e\} \tag{34}$$

where t_c is the corporate tax rate and $\varepsilon_e = \frac{de}{dt}\frac{1-t}{e}$ is the elasticity of income shifting from the income to corporate tax base with respect to the net-of-income tax rate.

When agents optimize perfectly and there are no private transfer costs, the three formulas are just different representations of the same equation, and should yield exactly the same estimate of marginal excess burden. The three formulas differ in the types of data that they employ. To implement Slemrod's and Saez's revenue-adjustment formulas, one must identify all the behavioral responses through which total tax revenue is affected following an income tax change. To implement the formula here, one must estimate the taxable and total income elasticities and the marginal resource cost of sheltering. In some applications, it may be easier to trace out revenue effects, while in others it may be easier to estimate the marginal resource cost of sheltering. Hence, the three formulas should be viewed as complements for empirical applications.

One benefit of the formula here is that it sheds light on the types of taxes and behavioral responses that generate the largest efficiency costs. In particular, it shows that the deadweight cost of sheltering is proportional to its resource cost, irrespective of its effects on other parts of the government's budget. In addition, only the formula proposed here accounts for transfers within the private sector and optimization errors in sheltering. This is because within-private-sector externalities and errors in perceived costs of sheltering do not show up on the government's budget. An instructive proof showing why the revenue-adjustment approaches cannot be applied in these situations is given in Appendix B.

III Conclusion

This paper has extended Feldstein's (1999) taxable income formula for deadweight loss to an environment in which agents' perceived private costs of evasion and avoidance differ from their social costs. The generalized formula shows that, to calculate the excess burden of a change in the income tax, one must determine (1) how much of the taxable income elasticity is driven by variation in labor supply choices vs. sheltering ($\varepsilon_{TI}, \varepsilon_{LI}$); and (2) the marginal resource cost of sheltering (μ). These factors are particularly important in understanding the efficiency cost of taxing individuals who have substantial ability to shelter income, such as high-income and self-employed individuals.

Characterizing the resource costs of reporting lower income is also important for topics beyond optimal income taxation. For example, Gruber and Joshua D. Rauh (2007) estimate the sensitivity of reported corporate profits to corporate tax rates to calculate the excess burden of the corporate tax. If corporations' reported profits are sensitive to tax rates primarily because of reporting effects and these changes in reporting do not have substantial resource costs, the excess burden from corporate taxation may be smaller than implied by Gruber and Rauh's calculations. Another example is Looney and Monica Singhal's (2006) use of the taxable income measure to estimate intertemporal substitution elasticities using anticipated changes in tax rates. Although individuals may shift their reported taxable income significantly across periods to minimize their tax burdens, labor supply and total earnings could be less elastic intertemporally (Goolsbee 2000). Since aggregate output is what matters in calibrating macroeconomic models, the resource costs of intertemporal shifting must be quantified in order to translate Looney and Singhal's estimates into the relevant intertemporal elasticity.

Estimating the marginal resource cost of sheltering, $g'(e)$, is especially important because its potential values span a very wide range. Many forms of sheltering appear to have low accounting costs relative to top marginal tax rates (approximately 40% in the U.S.). For instance, the accounting costs of conducting business in cash to under-report taxable income or setting up offshore accounts, trusts, and foundations are typically less than 5% of the amount sheltered. There could, however, be substantial economic resource costs from such sheltering behaviors, such as the need to maintain an inefficiently small firm to facilitate under-reporting or the loss of control inherent in delegating money to trusts and foundations. One must also account for resource costs that arise from distortions in consumption patterns induced by tax-sheltering motives, such as overconsumption of tax-favored goods (e.g. housing and healthcare) or business-related expenses (e.g. flying in first class, having a lavish office). Once these additional economic resource costs of sheltering are taken into account, $g'(e)$ could potentially be close to the marginal tax rate. Depending on where $g'(e)$ lies in the range from zero to the tax rate, the excess burden due to tax sheltering behavior could range from zero to the large values implied by existing studies of the elasticity of taxable income.

To estimate the marginal resource cost of sheltering, one must develop a mapping between the primitive parameter $g'(e)$ and observable behaviors (Chetty 2008). One promising approach to this problem is to compare the effects of tax changes on reported income and consumption bundles, recognizing that real resource costs ultimately distort consumption patterns (Yuriy Gorodnichenko, Jorge Martinez-Vazquez, and Klara Sabirianova Peter 2008).[13] Gorodnichenko et al. show that a large tax reform in 2001 in Russia induced substantial changes in reported taxable income but little change in the level and composition of consumption. Their findings suggest that the resource costs of changes in reported taxable income are small in the Russian case. Additional studies on the effects of tax reforms on consumption would be valuable given the prevalence of evasion and avoidance even in the U.S.: recent studies estimate that the evasion tax gap in the U.S. is 15% of tax revenue and that the avoidance tax gap is also substantial (Slemrod and Yitzhaki 2002, Slemrod 2007).

[13] The benefit of the consumption measure is that it automatically aggregates resource costs across different types of sheltering behaviors. Without this aggregate measure, one would have to calculate the marginal resource cost for each sheltering behavior individually and compute the weighted mean value to arrive at the relevant value of $g'(e)$, as in (32). The disadvantage of the consumption measure is that it does not capture non-monetary costs, such as the cost of violating ethical principles by evading taxes.

References

Allingham, Michael G., and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics,* 1(3-4): 323-338.

Andreoni, James, Brian Erard, and Jonathan Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature*, 36: 818-860.

Auerbach, Alan J. 1985. "The Theory of Excess Burden and Optimal Taxation." In Vol. 1, *Handbook of Public Economics*, ed. Alan J. Auerbach and Martin S. Feldstein, 67-127. Amsterdam: Elsevier Science Publishers.

Auerbach, Alan J., and James R. Hines Jr. 2002. "Taxation and economic efficiency." In Vol. 3, *Handbook of Public Economics*, ed. Alan J. Auerbach and Martin S. Feldstein, 1347-1421. Amsterdam: Elsevier Science Publishers.

Chetty, Raj. 2008. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." UC-Berkeley mimeo.

Chetty, Raj, Adam Looney, and Kory Kroft. 2008. "Salience and Taxation: Theory and Evidence." *American Economic Review.* Forthcoming.

Clotfelter, Charles T. 1983. "Tax Evasion and Tax Rates: An Analysis of Individual Returns." *Review of Economics and Statistics*, 65(3): 363-373.

Clotfelter, Charles T. 1985. *Federal Tax Policy and Charitable Giving.* Chicago: University of Chicago Press.

Dahlby, Bev. 1998. "Progressive taxation and the social marginal cost of public funds." *Journal of Public Economics,* 67(1): 105-122.

Feldstein, Martin S. 1995. "The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act." *Journal of Political Economy,* 103(3): 551-572.

Feldstein, Martin S. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics,* 81(4): 674-680.

Feldstein, Martin S. 2006. "The Effect of Taxes on Efficiency and Growth." *Tax Notes*, 111(6): 679-684.

Feldstein, Martin S., and Amy Taylor. 1976. "The Income Tax and Charitable Contributions." *Econometrica*, 44(6): 1201-1222.

Goolsbee, Austan. 1999. "Evidence on the High-Income Laffer Curve from Six Decades of Tax Reform." *Brookings Papers on Economic Activity*, 30(1999-2): 1-64.

Goolsbee, Austan. 2000. "What Happens When You Tax the Rich? Evidence from Executive Compensation." *Journal of Political Economy*, 108(2): 352-378.

Gordon, Roger H., and Joel B. Slemrod. 2002. "Are 'Real' Responses to Taxation Simply Shifting Between Corporate and Personal Tax Bases?" In *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, ed. Joel Slemrod. Cambridge: Harvard University Press.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter. 2008. "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion and Productivity Response in Russia." NBER Working Paper 13719.

Gruber, Jonathan, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics*, 84(1): 1-32.

Gruber, Jonathan, and Joshua D. Rauh. 2007. "How Elastic is the Corporate Income Tax Base?" In *Taxing Corporate Income in the 21st Century*, ed. Alan J. Auerbach, James R. Hines, and Joel B. Slemrod. Cambridge: Cambridge University Press.

Hessing, Dick J., Hank Elffers, Henry S.J. Robben, and Paul Webley. 1992. "Does deterrence deter? Measuring the effect of deterrence on tax compliance in field studies and experimental studies." In *Why people pay taxes: tax compliance and enforcement*, ed. Joel Slemrod. Ann Arbor: University of Michigan Press.

Joint Economic Committee, U.S. Congress. 2001. "Economic Benefits of Personal Income Tax Rate Reductions."

Kaplow, Louis. 1991. "Optimal Taxation with Costly Enforcement and Evasion." *Journal of Public Economics*, 43(2): 221-236.

Kaplow, Louis. 1995. "A Note on Subsidizing Gifts." *Journal of Public Economics* 58(3): 469-477.

Lindsey, Lawrence B. 1987. "Individual Taxpayer Response to Tax Cuts: 1982-1984, with Implications for the Revenue Maximizing Tax Rate." *Journal of Public Economics*, 33(2): 173-206.

Looney, Adam, and Monica Singhal. 2006. "The Effect of Anticipated Tax Changes on Intertemporal Labor Supply and the Realization of Taxable Income." NBER Working Paper 12417.

Mayshar, Joram. 1991. "Taxation with Costly Administration" *Scandinavian Journal of Economics*, 93(1): 75-88.

Pissarides, Christopher A., and Guglielmo Weber. 1989. "An Expenditure-Based Estimate of Britain's Black Economy." *Journal of Public Economics*, 39(1): 17–32.

Saez, Emmanuel. 2004. "Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications." In Vol. 18, *Tax Policy and the Economy*, ed. James Poterba. Cambridge: MIT Press.

Scholz, John T., and Neil Pinney. 1995. "Duty, Fear, and Tax Compliance: The Heuristic Basis of Citizenship Behavior." *American Journal of Political Science*, 39(2): 490-512.

Slemrod, Joel B. 1992. "Do Taxes Matter? Lessons from the 1980s." *American Economic Review*, 82(2): 250–256.

Slemrod, Joel B. 1995. "Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986." *American Economic Review*, 85(2): 175-180.

Slemrod, Joel B. 1998. "Methodological Issues in Measuring and Interpreting Taxable Income Elasticities." *National Tax Journal*, 51(4): 773-788.

Slemrod, Joel B. 2001. "A General Model of the Behavioral Response to Taxation." *International Tax and Public Finance*, 8: 119–128.

Slemrod, Joel B. 2007. "Cheating Ourselves: The Economics of Tax Evasion." *The Journal of Economic Perspectives*, 21(1): 25-48.

Slemrod, Joel B., and Shlomo Yitzhaki. 1996. "The Cost of Taxation and the Marginal Efficiency Cost of Funds." *International Monetary Fund Staff Papers*, 43(1): 172-198.

Slemrod, Joel B., and Shlomo Yitzhaki. 2002. "Tax avoidance, evasion, and administration." In Vol. 3, *Handbook of Public Economics*, ed. Alan J. Auerbach and Martin S. Feldstein, 1423-1470. Amsterdam: Elsevier Science Publishers.

Yitzhaki, Shlomo. 1974. "A Note on 'Income Tax Evasion: A Theoretical Analysis.'" *Journal* *of Public Economics*, 3(2): 201-202.

Yitzhaki, Shlomo. 1987. "On the Excess Burden of Tax Evasion." *Public Finance Review*, 15(2): 123-137.

Appendix A. Auditing Model with Risk Averse Agent and Administrative Costs

This appendix extends the analysis in section I.D in two ways. First, it models risk bearing as an excess burden of tax evasion, as in Yitzhaki (1987). Second, it shows how resource costs of auditing or fines can be taken into account.

Consider an economy populated by a set of identical agents of measure 1. Let each agent's utility be denoted by $u(c)$, which is increasing and concave. Agents are audited with probability $p(e)$, and the probability of audit independent across agents. In the state where an agent successfully evades taxes and is not audited, his income is $y + (1-t)(wl - e) + e$, where y represents unearned income. In the state where he is audited and caught, his income is $y + (1-t)wl - F(e)$ where y denotes unearned income. Suppose that the government bears a resource cost of auditing given by an increasing, convex function $\widetilde{K}_1(p(e))$ and a resource cost of imposing fines (e.g. running a prison) given by an increasing, convex function $\widetilde{K}_2(F(e))$. Let $K(e) = \widetilde{K}_1(p(e)) + \widetilde{K}_2(F(e))$ denote total administrative resource costs as a function of the amount of evasion.

Let $V(t, y)$ denote the agent's expected utility as a function of the tax rate and unearned income and $E(t, V)$ denote the corresponding expenditure function. Following Auerbach (1985), I define excess burden using the compensated variation measure:

$$EB = E(t, V(0, y)) - y - R(t) \tag{35}$$

where

$$R(t) = t(wl^c(t) - e^c(t)) + p(e^c(t))[te^c(t) + F(e^c(t))] - K(e^c(t))$$

is tax revenue net of administrative costs and $l^c(t)$ and $e^c(t)$ are income-compensated Hicksian demand functions. Given the continuum of agents, tax revenue is deterministic. To calculate excess burden, assume that tax revenue is returned lump-sum to every agent irrespective of whether he is audited or not.

Our objective is to derive an elasticity-based expression for marginal excess burden, $\frac{dEB}{dt}$. To begin, observe that the agent's indirect utility function is

$$V(t, y) = \max_{e,l}(1 - p(e))u(y + (1-t)(wl - e) + e) + p(e)u(y + (1-t)wl - F(e)) - \psi(l) \tag{36}$$

The expenditure function is

$$E(t, V) = \min_{e,l,\mu} y + \mu(V - (1-p(e))u(y+(1-t)(wl-e)+e) - p(e)u(y+(1-t)wl-F(e)) + \psi(l)) \tag{37}$$

where μ denotes the multiplier on the utility constraint. Let c_h denote consumption in the "good" state (where the agent is not caught) and c_l denote consumption is the state where he is caught and $Eu'(c) = (1-p(e))u'(c_h) + p(e)u'(c_l)$ denote the expected marginal utility of consumption. Using the first order conditions from agent optimization, it is easy to see that $\mu = \frac{1}{Eu'(c)}$. Note that $c_h - c_l = te + F$. Hence

$$\frac{dE}{dt} = \frac{1}{Eu'(c)}\{(1-p(e))u'(c_h)(wl - e) + p(e)u'(c_l)wl\} \tag{38}$$

$$\frac{dR}{dt} = wl - e + p(e)e + t\frac{d[wl^c]}{dt} + \frac{de^c}{dt}(-t + p'(e)[te + F] + p(e)[t + F'(e)] - K'(e)) \tag{39}$$

It follows that

$$-\frac{dEB}{dt} = t\frac{d[wl^c]}{dt} + p(e)(1-p(e))e\frac{[u'(c_h)-u'(c_l)]}{Eu'(c)} + \frac{de^c}{dt}[(-t+p'(e))(te+F)+p(e)(t+F'(e))-K'(e)] \tag{40}$$

To simplify the third term in this expression, consider the agent's first order condition with respect to the choice of e:

$$t = p'(e)\frac{[u(c_h)-u(c_l)]}{u'(c_h)} + p(e)\frac{[u'(c_h)t+u'(c_l)F'(e)]}{u'(c_h)} \tag{41}$$

Using this expression and collecting terms, we obtain

$$-\frac{dEB}{dt} = t\frac{d[wl^c]}{dt} + p(e)(1-p(e))e\frac{[u'(c_h)-u'(c_l)]}{Eu'(c)} + \frac{de^c}{dt}[p'(e)\{c_h - c_l - \frac{u(c_h)-u(c_l)}{u'(c_h)}\} + p(e)F'(e)\{\frac{u'(c_h)-u'(c_l)}{u'(c_h)}\} - K'(e)] \tag{42}$$

To clarify the intuition for this formula, I take a quadratic approximation to the utility function around c_h and write the formula in terms of the coefficient of relative risk aversion, $\gamma = \frac{-u''(c)}{u'(c)}c$:

$$-\frac{dEB}{dt} \approx t\frac{d[wl^c]}{dt} - \gamma\{p(e)(1-p(e))e\frac{\Delta c}{c}\} - \gamma\frac{de^c}{dt}\{p(e)F'(e)\frac{\Delta c}{c} + \frac{1}{2}p'(e)c_h(\frac{\Delta c}{c})^2 - K'(e)\} \tag{43}$$

where $\frac{\Delta c}{c} = \frac{te+F}{c_h}$ denotes the percentage loss in private income when the agent is caught. Note that when $\gamma = 0$ (risk neutrality), the last two terms drop out of (43) and it coincides with (11) as expected. The second term reflects the cost of the additional risk directly from the increase in the tax rate, which raises the difference in income between the two states in proportion to the amount of tax evaded e. The third term reflects the cost of the additional risk that the agent bears from one dollar of additional evasion, due to the increased fine and probability of audit. Both of these additional risk costs constitute real resource costs because they reduce net social welfare.

Define the marginal resource cost of evasion as $g'(e) = \{p(e)F'(e)\frac{\Delta c}{c} + \frac{1}{2}p'(e)c_h(\frac{\Delta c}{c})^2 - K'(e)\}$ and the marginal resource cost directly from the increase in the tax rate as $g'_t(e) = \gamma\{p(e)(1-p(e))e\frac{\Delta c}{c}\}$. Then we can rewrite (43) as

$$\frac{dW}{dt} = -\frac{dEB}{dt} = t\frac{d[wl^c]}{dt} - g'(e)\frac{de^c}{dt} - g'_t(e) \tag{44}$$

$$= -\frac{t}{1-t}\{\mu\varepsilon^c_{TI}(wl-e) + (1-\mu)\varepsilon^c_{LI}wl\} - g'_t(e) \tag{45}$$

where ε^c_{TI} denotes the compensated taxable income elasticity and ε^c_{LI} denotes the compensated total earned income elasticity. This equation coincides with the formula for the general case with resource and transfer costs in (18) except for the additional term $g'_t(e)$ reflecting the direct cost of subjecting the agent to more risk when t is increased. Excess burden still depends on a weighted average of the (compensated) taxable income and total earned income elasticities. The weight is still proportional to $g'(e)$, which is now redefined to include the risk-bearing costs of evasion and administrative costs paid by the government.

Appendix B. Necessary Conditions for Revenue-Adjustment Formulas

In the general case with resource costs, transfer costs, and optimization errors, recall from section I.E that

$$\frac{dW}{dt} = t\frac{dLI}{dt} - g'(e)\frac{de}{dt} \tag{46}$$

To understand the correspondence between this formula and the revenue-adjustment approaches of Slemrod (1998) and Saez (2004), it is useful to distinguish between three cases.

1. Revenue offset with agent optimization. Suppose $z(e)$ accrues to the government, so that $R = t(wl - e) + z(e)$. Then $\frac{\partial R}{\partial t}|_{l,e} = wl - e$ while $\frac{dR}{dt} = wl - e + t\frac{d(wl-e)}{dt} + \frac{\partial z}{\partial e}\frac{de}{dt}$. If the agent optimizes his sheltering decision, $g'(e) = t - z'(e)$. Hence the standard tax revenue adjustment formula holds:

$$\frac{dW}{dt} = \frac{dR}{dt} - \frac{\partial R}{\partial t}|_{l,e} \tag{47}$$

2. Private transfer cost with agent optimization. Suppose $z(e)$ is a private transfer to another agent in the economy. Then $\frac{dR}{dt} = wl - e + t\frac{d(wl-e)}{dt}$, and we obtain

$$\frac{dW}{dt} = \frac{dR}{dt} - \frac{\partial R}{\partial t}|_{l,e} + z'(e)\frac{de}{dt} \tag{48}$$

where the additional term reflects the externality transfer within the private sector that occurs through the agent's behavioral response to the tax change. The revenue adjustment formula does not hold with private transfers.

3. Revenue offset with optimization error in sheltering. As in case 1, $z(e)$ accrues to the government, so that $R = t(wl - e) + z(e)$, and $\frac{dR}{dt} = wl - e + t\frac{d(wl-e)}{dt} + \frac{\partial z}{\partial e}\frac{de}{dt}$. But we cannot equate $g'(e)$ with $t - z'(e)$ because the agent's first-order-condition for e does not hold. Hence the simplest revenue-adjustment formula that can be obtained is

$$\frac{dW}{dt} = \frac{dR}{dt} - \frac{\partial R}{\partial t}|_{l,e} + (t - z'(e) - g'(e))\frac{de}{dt} \tag{49}$$

where the additional term reflects the gap between the tax rate and the sum of the transfer and resource costs caused by the agent's optimization error. In the case where the agent optimizes, $t - z'(e) - g'(e) = 0$, and the formula collapses to the standard revenue-adjustment formula.

These derivations establish that the revenue adjustment formula holds if and only if all transfer costs are paid to the government and agents optimize their sheltering decisions. The Saez formula in (34) is an alternative representation of the revenue adjustment formula in (33) for the special case with two tax bases, and hence this result applies to that formula as well.