Marginal Deadweight Loss with Nonlinear Budget Sets

Sören Blomquist and Laurent Simula
Uppsala University and Uppsala Center for Fiscal Studies*†

May 11, 2015

Abstract

Most theoretical work on how to calculate the marginal deadweight loss has been done for linear taxes and for variations in linear budget constraints. However, most income tax systems are nonlinear, generating nonlinear budget sets. A usual procedure to calculate the marginal deadweight loss for variations in nonlinear income taxes has been to linearize the nonlinear budget constraint. Our main theoretical result is that the overall curvature of the tax system plays is as important as the curvature of indifference curves when computing deadweight loss measures. Using numerical simulations calibrated on US data, we then show that the apparently innocuous linearization may lead to a substantially large overestimation of the marginal deadweight loss.

Keywords: Deadweight Loss, Taxable Income, Nonlinear Budget Constraint
JEL classification: H21, H24, H31, D61

1 Introduction

There is a large literature on the excess burden of taxation (see Auerbach and Hines (2002) for an excellent survey). A first strand of the literature is devoted to the comparison of different measures of this burden, and looks at variations in linear prices caused by commodity taxation or a linear wage tax. A second strand of the literature addresses the issue of redistribution through progressive taxation. The definition of the excess burden is then even more difficult because there are many ways of varying a nonlinear

*Uppsala Center for Fiscal Studies, Uppsala University, Department of Economics, P.O. Box 513, SE-75120 Uppsala,Sweden. E-mail: soren.blomquist@nek.uu.se and laurent.simula@nek.uu.se

†We are grateful to Spencer Bastani, Thomas Gaube, Laurence Jacquet, Henrik Kleven, Klaus Kreiner, Jean-Marie Lozachmeur, Olof Johansson Stenman, Etienne Lehmann, Luca Micheletto, Håkan Selin and Alain Trannoy for very helpful comments and suggestions.

tax. In addition, redistribution only makes sense if agents are heterogeneous, which leads to the tricky question of how individual effects of taxation should be aggregated. These serious problems have been very carefully investigated, but a common procedure has been to locally linearize the budget constraint instead of fully accounting for its nonlinearity. The present article does not contribute on the alternative ways of defining the excess burden of taxation or to aggregate individual gains and losses. Taking a given measure and aggregation procedure, we present a method for properly determining the marginal deadweight loss and marginal tax revenue for a *nonlinear* income tax. We contrast our method with the linearization approach, explaining how the latter overstates marginal deadweight losses and underestimates marginal revenues (and thus multiplicatively overstates the ratio) when marginal tax rates are increasing with the income level. We further provide numerical simulations that apply the method and show the extent of the bias, which is large as a pratical matter.

It is simpler to illustrate the basic ideas under the assumption that the budget constraints are smooth and convex. The utility curvature tells us that, when a person starts adjusting (e.g., labor supply), the marginal utility is changing rapidly; hence little adjustment is needed to restore the first-order conditions. Now, when the budget set has curvature, that curvature is in essence doing the same: as the person adjusts, the price is changing more rapidly, so one restores the first-order conditions more quickly, i.e, for a smaller behavioral change. Consequently, if the budget constraint is linearized, one sets the curvature to zero and overestimates the change in taxable income. The problem when linearizing is therefore that one misses the fact that the curvature of the budget constraint and the curvature of the indifference curve are of *equal* importance for how large a change in taxable income will be.

Actual tax systems are usually piecewise linear. We show that, in this case, deadweight loss measures critically depend on the overall curvature of the tax system.[1]. Using numerical simulations calibrated to reflect to US economy in 1979, 1993 and 2006, we find that the apparently innocuous linearization may lead to substantially large measurement errors in both the marginal deadweight loss and the marginal tax revenue. These errors are magnified when one calculates the marginal deadweight loss per marginal tax dollar. For a low elasticity of taxable income (0.2), we find that the linearization overestimates the latter by 4.1% in 2006, 7.2% and 9.3% in 1979. For larger but yet reasonable elasticities (0.4), the bias is of 10.7% in 2006, 25.3% in 1993 and 24.5% in 1979. However,

[1]This is due to different errors and optimization frictions which should be accounted for when agents make their optimal choices. See, e.g., Burtless and Hausman (1978), Hausman (1979), Blomquist (1983) and Hausman (1985)) and more recently Chetty (2012).

all these errors were much larger in 1979 and 1993 than in 2006. This is due to the fact that the overall curvature of the US tax system was much less pronounced in 2006 than a few decades ago. This evolution has been observed in most developed countries, until the financial crisis of 2007. In the afternath of the crisis, the marginal tax rates faced by high income earners have been raised in most countries, and further increases are vividly discussed. In this context, we believe that it is critical to correctly account for the curvature of the tax function to correctly assess the impact of potential tax reforms.

The article is organized as follows. In section 3 and 4, we use smooth budget constraints to introduce the main idea and show how the marginal deadweight loss and marginal tax revenue loss should be calculated. Section 3 introduces the basic definitions and Section 4 presents our main theoretical findings. We show that the curvature of the budget constraint is as important as the curvature of the indifference curves when computing deadweight loss measures. We also give expressions for how large the bias using a linearization procedure can be. In Section 5, we present calculations for the US tax system for three different years and numerically establish that the apparently innocuous linearization procedure may lead to large estimation errors. Section 6 concludes.

2 Review of the Literature

The study of the deadweight loss (or excess burden) of taxation has a long tradition in economics going back as far as Dupuit (1844). Modern type of empirical work on the deadweight loss of taxation is heavily influenced by the important work of Harberger in the 1950s and 1960s (see for example Harberger (1962, 1964)). A second generation of empirical work was inspired by Feldstein (1995, 1999). Feldstein argued that previous studies had neglected many important margins that are distorted by taxes. By estimating how total taxable income reacts to changes in the marginal tax, one would be able to capture distortions of all relevant margins. Feldstein's own estimates indicated large welfare losses whereas many later studies arrived at estimates of the welfare loss that were larger than those obtained in pre-Feldstein studies, but considerably lower than the estimates obtained by Feldstein. An important ingredient in modern studies of the deadweight loss of taxes is the estimation of a (Hicksian) taxable income supply function(Gruber and Saez, 2002; Kopczuk, 2005; Saez, Slemrod, and Giertz, 2012). These taxable income functions show how taxable income varies as the slope of a linear budget constraint of individuals is changed at the margin.

Kleven and Kreiner (2006), Eissa, Kleven, and Kreiner (2008) and Gelber and Mitchell (2012) are recent important articles that employ a linearized budget constraint to calculate deadweight losses. Kleven and Kreiner (2006) extend the theory and measurement of the marginal cost of public funds to account for labour force participation responses. Eissa, Kleven, and Kreiner (2008) embed the participation margin in an explicit welfare theoretical framework and show that not modelling the participation decision induce large errors. Gelber and Mitchell (2012) examine how income taxes affect time allocation during the entire day, and how these time allocation decisions interact with expenditure patterns.

In empirical work estimating behavioral parameters, it is also common to linearize budget constraints and pursue the analysis with the linearized budget constraints. However, when doing this, one is well aware that this linearization creates econometric problems and there is a large literature on methods to handle these problems (see e.g., Burtless and Hausman (1978), Hausman (1985) and Blomquist and Newey (2002)).

There are very few studies addressing the problems that arise when one linearizes budget constraints and use the linearized budget constraints to do welfare analysis. In an important article, Dahlby (1998) describes the calculation of the social marginal cost of public funds for a progressive tax system which distorts individuals' labour supply decisions. He first notes that a single-person model, as the one used by most of the theoretical literature on the marginal cost of public funds, is inappropriate for a progressive tax system. He therefore investigates the case with heterogeneous individuals. He develops a multi-person framework and derives an expression for the social marginal cost of public funds for a generic piecewise linear tax system. He then considers more specific tax reforms and shows that several well-known formulas are obtained as special cases. We employ a multi-person framework as in Dahlby (1998). However, contrary to the latter, our analysis starts from a smooth tax system, in which case the curvature of the budget constraints is shown to play the same part as the curvature of the indifference curves. Introducing errors and optimization frictions, we see that even though the statutory tax system is piecewise linear, its overall curvature plays a key part when computing deadweight loss measures.

3 Basic Setting

Statutory tax systems are usually piecewise linear. However, in order to get simple and clean results, we start our analysis by considering smooth budget constraints. The main reason is that it is basically the *general shape* of the tax system and the budget constraints that determine the average behavior. In the next section, we will consider a piecewise linear tax schedule and add errors and optimization frictions. We will see that the agents behave as if the tax system were smooth even though the *statutory* tax system consists of several linear segments.

3.1 The Tax System

A *linear* income tax can be varied in two ways. One can change the intercept, which leads to a pure income effect, or change the proportional tax rate, which leads to a substitution and an income effect. For a *nonlinear* income tax, there are many more possible ways to vary the tax. Break points can be changed, the intercept can be changed and the slope can also be changed. Moreover, the slope can be changed in different ways. We do not cover all these different possibilities to vary a nonlinear tax. We focus on a particular kind of change in the slope, namely a change in the slope such that the marginal tax changes with the same number of percentage points at all income levels. The general insight of the article that the comparative statics for the compensated taxable income depends on the curvature of the budget constraint carries over to variations in other tax parameters. We could write the tax system as $T(A) = g(A, \gamma)$, where γ is a tax parameter the variation of which is under scrutiny. Then, it would be the curvature of g, $\partial^2 g/\partial A^2$, that would be critical for our results. What we do below is a special case, which we believe makes the analysis more transparent and directly applicable.

We model the tax in the following way. Let A denote taxable income and the tax on A be given by $T(A)$. The results below depend on the curvature of the tax function, $\partial^2 T(A)/\partial A^2$. For simplicity, we show details for a specific formulation $T(A) = g(A) + tA$, with $g'(A) > 0$, $g''(A) > 0$ and $t \geq 0$. Note that in this case $\partial^2 T(A)/\partial A^2$ reduces to $g''(A)$. For example, we can think of $g(A)$ as a nonlinear federal tax. There are several alternative interpretations of tA. It could be a payroll tax, a value added tax or a proportional state income tax. Within the Scandinavian framework, it could be interpreted as the local community income tax. What we study is the marginal deadweight loss of an increase in t. A change in t implies that the marginal tax is increased by the same number of percentage points at all income levels.

There are two good reasons why we have parameterized the tax system in the way described above. When we vary the slope of a linear budget constraint, the intercept will not change. Hence, a change in t will not change the virtual income but only the slope, thereby giving an experiment similar to a change in the slope of a linear budget constraint.[2] A second reason is, of course, that $g(A) + tA$ is a good approximation of actual tax systems.

[2] This nice feature of the parameterization used was pointed out to us by Håkan Selin.

3.2 The Marginal Deadweight Loss

Consider the utility maximization problem:

$$\max_{C,A} U(C, A, v) \text{ s.t. } C \leq A - g(A) - tA + B, \tag{P1}$$

where C is consumption, v an individual specific preference parameter and B lump-sum income. We assume that the utility function $U(C, A, v)$ has the usual properties. We call $C(t, B, v)$ and $A(t, B, v)$ the solutions to problem (P1). The form of these two functions depends on the functional forms of U and g. Sticking $C(t, B, v)$ and $A(t, B, v)$ back into the utility function, we obtain the indirect utility $\overline{u}(v) := U(C(t, B, v), A(t, B, v), v)$. For each individual, the latter is the maximum utility level obtained under the given tax system. Because individuals have different v's, they choose different taxable incomes and have different $\overline{u}(v)$.

We now study the marginal deadweight loss of a small increase in the tax parameter t. We also examine the marginal deadweight loss per marginal tax dollar, as this measure is often used in the literature. We first derive expressions without resorting to any linearization. For this purpose, we define the expenditure function as:

$$E(t, v, \overline{u}) = \min_{C,A} \{C - A + g(A) + tA - B\} \text{ s.t. } U(C, A, v) \geq \overline{u}. \tag{P2}$$

This problem also defines the compensated demand and supply functions, $C^h(t, v, \overline{u})$ and $A^h(t, v, \overline{u})$ respectively, where the superscript h denotes that it is Hicksian functions. It is important to note that these functions depend on the functional form of $U(C, A, v)$ *and on the functional form of* $g(A)$. In almost all empirical and theoretical analyses, we work with demand and supply functions generated by linear budget constraints. In contrast, the functions defined by (P1) and (P2) are generated by a nonlinear budget constraint. In addition, we define the compensated tax revenue function as:

$$R(A^h(t, v, \overline{u})) = g(A^h(t, v, \overline{u})) + tA^h(t, v, \overline{u}) \tag{1}$$

and the marginal tax revenue – whilst keeping utility constant – as:

$$MTR := \frac{dR(A^h)}{dt} = A^h + (g'(A^h) + t)\frac{dA^h}{dt}. \tag{2}$$

The marginal deadweight loss is often measured as the difference between compensated changes in expenditure and collected taxes, i.e.,

$$\begin{aligned} MDW &:= \frac{dE\left(t,v,\overline{u}\right)}{dt} - \frac{dR(A^h\left(t,v,\overline{u}\right))}{dt} \\ &= A^h - g'(A^h)\frac{dA^h}{dt} - A^h - t\frac{dA^h}{dt} = -(g'\left(A^h\right) + t)\frac{dA^h}{dt}, \end{aligned} \tag{3}$$

where we used the envelope theorem to obtain $dE\left(t,v,\overline{u}\right)/dt = A^h$. We will use this definition of the marginal deadweight loss in this article. We know that there are alternative definitions in the literature, but our main point is independent of this aspect. The ratio (3)/(2) corresponds to the marginal deadweight loss per marginal tax dollar.

3.3 Linearization Procedure

When calculating the marginal deadweight loss for a non-linear tax system, it is common to linearize the budget constraint and proceed as if the budget constraint were linear. This linearization is done in the following way. Let us consider particular values v^*, t^* and B^* and the solution to (P1), $C^* = C\left(t^*,v^*,B^*\right)$ and $A^* = A\left(t^*,v^*,B^*\right)$. We can linearize the budget constraint around this point with local prices defined by $p_c = 1$ and $p_A = g'\left(A^*\right) + t^*$ to obtain the linear budget constraint $C = A - p_A A + M$, where M is defined as $M = C^* - A^* + p_A A^*$. We now consider the problem:

$$\max_{C,A} U\left(C,A,v^*\right) \text{ s.t. } C \leqq A - p_A A + M. \tag{P3}$$

We call $C_L\left(p_A,v^*,M\right), A_L\left(p_A,v^*,M\right)$ the solution to this problem. Here, we use the subscript L to show that these are functions generated by a linear budget constraint.

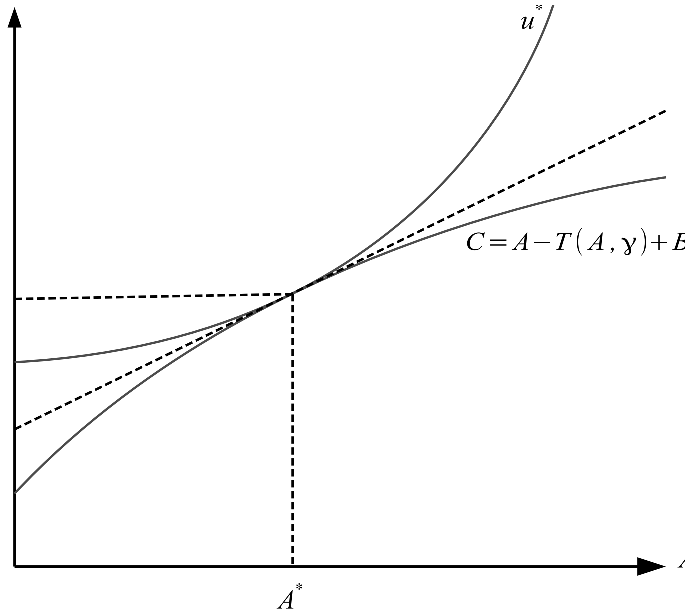
We define the expenditure function corresponding to this linear budget constraint as

$$E_L\left(t,v,\overline{u}\right) = \min_{C,A}\left\{C - A + p_A A - M\right\} \text{ s.t. } U\left(C,A,v\right) \geq \overline{u} \tag{P4}$$

and denote its solution by $C_L^h\left(t,v,\overline{u}\right), A_L^h\left(t,v,\overline{u}\right)$, where the subscript L indicates that it is the solution to a problem where the objective function is linear and the superscript h that this is Hicksian demand-supply functions. Let us define the compensated revenue function as: $R\left(A_L^h\left(t,v,\overline{u}\right)\right) = g\left(A_L^h\left(t,v,\overline{u}\right)\right) + tA_L^h\left(t,v,\overline{u}\right)$. Marginal tax revenue, keeping utility constant, is given by:

$$MTR_L := \frac{dR(A_L^h)}{dt} = A_L^h + (g'(A_L^h) + t)\frac{dA_L^h}{dt}. \tag{4}$$

Figure 1: Nonlinear and linearized programmes



We obtain the marginal deadweight loss as:

$$MDW_L := \frac{dE_L(t, v, \overline{u})}{dt} - \frac{dR_L(A_L^h(t, v, \overline{u}))}{dt}$$
$$= A_L^h - g'(A_L^h)\frac{dA_L^h}{dt} - A_L^h - t\frac{dA_L^h}{dt} = -(g'\left(A_L^h\right) + t)\frac{dA_L^h}{dt}. \tag{5}$$

The marginal deadweight loss per marginal tax dollar is the ratio (5)/(4).

Figure 1 illustrates the links between the four problems that we have studied. The optimization problem (P1) maximizes utility given the curved budget constraint $C = A - g(A) - tA + B$. Let us consider particular values for the proportional tax and lump-sum income: t^* and B^*. Suppressing the dependence on v, we denote the solution by $C^* = C(t^*, B^*), A^* = A(t^*, B^*)$. This defines the utility level $u^* = U(C^*, A^*)$. Optimization problem (P2) minimizes expenditures to reach the utility level u^* for the given nonlinear tax system. By construction, the solution to this problem is also C^*, A^*. Linearizing around (A^*, C^*), so that the linear budget constraint is tangent to the indifference curve at (A^*, C^*), we have two other optimization problems. Problem (P3) maximizes utility subject to the linear budget constraint going through (A^*, C^*) and having the same slope as the indifference curve through (A^*, C^*). Problem (P4) is to minimize expenditures given the utility level u^* and the general shape of the budget constraint given by the linear budget constraint. By construction, the four optimization problems have the same solution. For any t and B, we thus have the identities

$$\begin{aligned} & A(t,B) \equiv A^h\left(U\left(C\left(t,B\right),A\left(t,B\right)\right)\right) \\ \equiv\ & A_L\left(p_A\left(C\left(t,B\right),A\left(t,B\right)\right),M\left(C\left(t,B\right),A\left(t,B\right)\right)\right) \equiv A_L^h\left(U\left(C\left(t,B\right),A\left(t,B\right)\right)\right). \end{aligned} \tag{6}$$

4 Main Result

In this section, we show that the curvature of the tax system is as important as the curvature of the indifference curves when computing the marginal deadweight loss. When we look at the definitions introduced above, expressions (3) and (5) look quite similar, as do expressions (2) and (4). By construction, it is true that $A_L^h = A^h$, implying that $g'\left(A_L^h\right) + t = g'\left(A^h\right) + t$. However, dA^h/dt and dA_L^h/dt usually differ, implying a bias when the linearization procedure is used. To show this, we start with a simple example, which we then generalize.

4.1 A Simple Example

To simplify notation, we in this example suppress the preference parameter v. We assume that the utility function takes the quasilinear form $U = C - \alpha A - \gamma A^2$. This implies that the income effect for the supply of A is zero, so that the Marshallian and Hicksian supply functions are the same. We assume that the tax is given by $T(A) = tA + pA + \pi A^2$, where we can interpret tA as the state tax and $pA + \pi A^2$ as the federal tax. This yields a budget constraint $C = A - (p+t)A - \pi A^2 + B$, where B is lump-sum income. Substituting the budget constraint into the utility function, we obtain $U = A - (p+t)A - \pi A^2 + B - \alpha A - \gamma A^2$. Maximizing with respect to A, we get $dU/dA = 1 - (p+t) - 2\pi A - \alpha - 2\gamma A$. We see that a necessary condition for a nonnegative A is $1 - (p+t) - \alpha \geq 0$. We find that $d^2U/dA^2 = -2(\pi+\gamma) < 0$ for $\pi + \gamma > 0$. Setting $dU/dA = 0$ and solving for A, we obtain

$$A = \frac{1 - (p+t) - \alpha}{2(\pi+\gamma)}. \tag{7}$$

Since we have the quasilinear form, this is also the Hicksian supply. We immediately have

$$\frac{dA^h}{dt} = -\frac{1}{2(\pi+\gamma)}. \tag{8}$$

From (8), we see that the size of the substitution effect depends on the curvatures of the indifference curve and the budget constraint. We note that it is immaterial whether the curvature emanates from the indifference curve or from the budget constraint. What matters is the curvature of the indifference curve in relation to the budget constraint. The larger the total curvature, given by $2(\pi + \gamma)$ in our example, the smaller is the change in taxable income and the smaller is the deadweight loss.

Suppose that we have particular values for the parameters of the problem and denote the solution (C^*, A^*). We can linearize the budget constraint around this point and get the budget constraint $C = A - [(p+t) + 2\pi A^*] A + M$, where $M = C^* - [1 - (p+t) - 2\pi A^*] A^*$. Given this linearization, an individual solves:

$$\max_{C,A} \left\{ C - \alpha A - \gamma A^2 \right\} \text{ s.t. } C \leqq A - [(p+t) + 2\pi A^*] A + M. \tag{9}$$

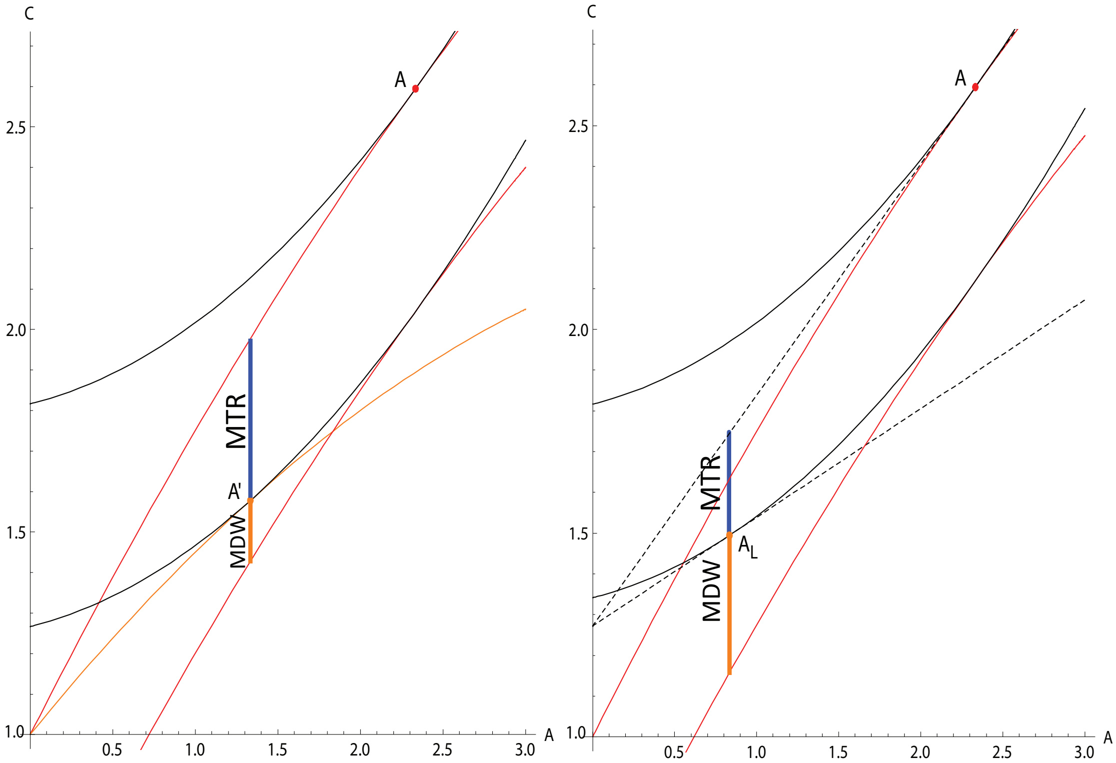
Substituting the binding budget constraint into the utility function, we want to maximize $A - [(p+t) + 2\pi A^*] A + M - \alpha A - \gamma A^2$. Denoting this expression by $\widetilde{U}$, we obtain $d\widetilde{U}/dA = 1 - (p+t) - 2\pi A^* - \alpha - 2\gamma A$ and $d^2\widetilde{U}/dA^2 = -2\gamma$. The second-order condition is satisfied for $\gamma > 0$. Setting $d\widetilde{U}/dA = 0$ and solving for A, we get $A_L^h = [1 - (p+t) - 2\pi A^* - \alpha] / (2\gamma)$ and

$$\frac{dA_L^h}{dt} = -\frac{1}{2\gamma}. \tag{10}$$

We see that a marginal increase in the tax rate t induces a smaller response in taxable income when the budget constraint is linearized. Given the definitions introduced in Section 3, this implies that the linearization procedure overestimates the marginal deadweight loss and underestimates the marginal tax revenue. To get an order of magnitude, suppose for example that $\pi = \gamma = 0.1$. We then have $dA^h/dt = -2.5$ while using the supply function generated by the linearized budget constraint gives $dA_L^h/dt = -5$. This means that the linearization procedure overestimates the deadweight loss with a factor 2. If we choose $\pi = 0.05$ and set all other parameters equal to 0.1, then the linearization procedure overestimates the correct deadweight with a factor 1.5 (1.5 instead of 1.0), underestimates the correct marginal tax revenue with a factor 1.6 (0.83 instead of 1.33) and overestimates the marginal deadweight loss per marginal tax dollar with a factor 2.4 (1.8 instead of 0.75).

In Figure 2, we illustrate the deadweight loss of a discrete change in t, from $t = 0$ to $t = 0.3$, for parameter values of $\alpha = \gamma = 0.1$, $p = 0.2$, $\pi = 0.05$ and $B = 1$. In the left panel, we show the correct calculation using a variation in the nonlinear budget

Figure 2: Deadweight loss when the budget constraint is nonlinear (left panel) and linearized (right panel)



constraint. The bundle chosen prior to the tax change is A, at the tangency point between the budget constraint and the highest feasible indifference curve. The increase in t shifts the nonlinear budget constraint in such a way that A' is now chosen instead of A. The deadweight loss corresponds to the difference between the equivalent variation and the variation in tax revenue, labelled MTR. It is thus shown by the thick vertical line MDW below A'. In the right panel, we show the standard procedure which employs a variation in the linearized budget constraint. The nonlinear budget constraint through A is linearized around this point. The increase in t induces a rotation of the linearized budget constraint around the intercept. The bundle A_L is now chosen instead of A. We see that the deadweight loss, shown by the thick vertical MDW line below A_L, is much larger than when the correct procedure is used. The change in tax revenue is given by the line MTR. Regarding the deadweight loss per tax dollar (given by the ratio of MDW and MTR), we see that the figure obtained when linearizing is very different from the correct one. Hence, the error made for the change in the deadweight loss and the error made for the change in tax revenue are magnified when one calculates the marginal deadweight loss per marginal tax dollar.

4.2 Generalization of the Example

We can easily generalize the example above. Let us consider the general utility function $U(C, A, v)$. The Hicksian supply function for taxable income is defined by problem (P2). We will reformulate this problem. The constraint $U(C, A, v) \geq \overline{u}$ is binding at the optimum and can thus be rewritten as $C = f(A, v, \overline{u})$, where the function f is defined by $U(f(A, v, \overline{u}), A, v) = \overline{u}$. Substituting the constraint $C = f(A, v, \overline{u})$ into the objective function, we obtain the minimization problem $\min_A f(A, v, \overline{u}) - A + tA + g(A) - B$. Let us for convenience use the notation $f'(\cdot)$ to denote $\partial f / \partial A$. The first order condition $f'(A, v, \overline{u}) - 1 + t + g'(A) = 0$ defines the Hicksian supply function $A^h(t, v, \overline{u})$. Differentiating it implicitly yields:

$$\frac{dA^h}{dt} = -\frac{1}{g'' + f''}. \tag{11}$$

In the analysis above, $f'(A, v, \overline{u})$ is the slope of the indifference curve. Hence, $f''(A, v, \overline{u})$ shows how the slope of the indifference curve changes as A is increased along the indifference curve and, thus, gives the curvature of the indifference curve. For the special case of a quasilinear utility function, with zero income effects for the taxable income function, $\overline{u}$ would not be an argument in the $f(\cdot)$ function. From (11), we see that the curvature of the budget constraint is as important for the size of the marginal deadweight loss as is the curvature of the indifference curve. *What matters is the curvature of the indifference curve in relation to the budget constraint.*[3] When the budget constraint is linear and $g'' = 0$, dA^h/dt reduces to $dA^h/dt = -1/f''$. Hence, if we linearized, we would obtain:

$$\frac{dA_L^h}{dt} = -\frac{1}{f''}, \tag{12}$$

which confirms that the linearization procedure leads to an overestimation of the true marginal deadweight loss.

In empirical studies of the taxable income function, it is usually the taxable income function $A_L^h(t, v, \overline{u})$, valid for a linear budget constraint, that is estimated and reported. However, if we know dA_L^h/dt as well as the tax function $T(A) = g(A) + tA$, it is easy

[3] In a recent article on optimal income taxation, ReP (2010) define elasticities along the nonlinear tax schedule, instead of the usual elasticities defined along linearized schedules. They show that the formula obtained in Saez (2001) can then be written in a more transparent way, which do not make use of the concept of virtual densities.

to calculate the comparative statics for the taxable income function $A^h(t, v, \overline{u})$. This is because the comparative statics for the two functions are related according to the formula:

$$\frac{dA^h}{dt} = \frac{dA_L^h/dt}{1 - g''(A)\left(dA_L^h/dt\right)}. \tag{13}$$

4.3 Bias when Linearizing

We below will measure the bias implied by linearizing in three ways: the overestimation of the marginal deadweight loss, the underestimation of the marginal tax revenue and overestimation of the marginal deadweight loss per marginal tax dollar. We will see that all these measures depend on the relative sizes of g'' and f''. For simplicity, we use a to denote the ratio g''/f''. Then, a is a measure of the relative curvature of the budget constraint and the indifference curve.

The relative error in the marginal deadweight loss when using the linearized budget constraint is given by the ratio of expressions (5) to (3), i.e. by:

$$\frac{MDW_L}{MDW} = \frac{dA_L^h/dt}{dA^h/dt} = \frac{g'' + f''}{f''} = 1 + \frac{g''}{f''} = 1 + a. \tag{14}$$

Hence, a is also a direct measure of the relative bias in the marginal deadweight loss if we incorrectly linearize. For example, if $a = 1$ and hence $g'' = f''$, the linearization procedure overstates the true marginal deadweight loss by a factor 2. *This holds true irrespective of the absolute size of g'' and f''. It is the relative curvature of the budget constraint and the indifference curve that matters.* Note that the bias a would be negative if the tax function were concave ($g''(A) < 0$). In this case, the linearization procedure would underestimate the correct marginal deadweight loss. In the numerical simulations provided in the next section, we fully account for nonconvexities of the budget set.

The relative curvature of the budget constraint and the indifference curve as measured by a also play an important role in the expression for the relative error in marginal tax revenue. However, f'' and g'' enter this expression in other ways. The relative error in the marginal tax revenue is obtained as the ratio of (4) and (2). Using (6) and (14), we get:

$$\frac{MTR_L}{MTR} = 1 + a \times \frac{(g'(A^h) + t)\ dA^h/dt}{A^h + (g'(A^h) + t)\ dA^h/dt}. \tag{15}$$

4.4 Marginal Deadweight Loss for the Population

If we want to find the aggregate marginal deadweight loss, we can integrate over v. From a welfare point of view, there is no obvious way how one should aggregate the marginal deadweight loss for different individuals. However, it is fairly common to calculate a weighted sum (or average) of the marginal deadweight losses of the different agents in the population. Whatever the weights that are used, it is clear that the aggregate marginal deadweight loss calculated with the function A_L^h gives a higher value than if calculated using A^h. In the next section, we will simply compute the average marginal deadweight loss

There is no clear way of aggregating the marginal deadweight loss per marginal tax dollar. For example, if one calculates the arithmetic average of the marginal deadweight loss per marginal tax dollar, individuals for which the marginal tax revenue is low would receive a very large weight in the aggregation process, which may be difficult to justify and misleading in terms of policy recommendations. For this reason, we in the next section will compute the marginal deadweight loss per tax dollar as the ratio between the average marginal deadweight loss and the average marginal tax revenue.

5 Numerical Illustration

We already know that the curvature of the budget constraint is as important as the curvature of the indifference curves when computing the marginal deadweight loss. The aim of this section is to assess the potential magnitude of the bias induced by the linearization procedure.

5.1 Random Shocks

Statutory tax systems are usually piecewise linear. However, because of errors and optimization frictions, individuals may behave as if the budget constraint was actually smooth.[4] To make this point, we consider a piecewise linear tax system $T(A)$ and introduce a random shock δ corresponding to an increase or decrease in taxable income. This shock has support $\left[\underline{\delta}, \overline{\delta}\right]$, pdf D, and mean equal to zero. Individuals know how it is distributed and behave as expected utility maximizers. Because of the latter, the former may actually behave as if the tax system was smooth even though it consists of a combination of linear brackets.

[4] The importance of random shocks and optimization frictions was already emphasized by Burtless and Hausman (1978), Hausman (1979), Blomquist (1983), Hausman (1985) and more recently Chetty (2012)).

Following **?**, we further assume that there is no income effect on taxable income. The Bernoulli utility is given by:

$$U = A + \delta - T(A+\delta) - \eta\left(\frac{A}{v}\right). \tag{16}$$

The heterogeneity parameter v corresponds to the individual ability to transform effort into earnings. The function η captures the disutility of effort, and satisfies $\eta' > 0$ and $\eta'' > 0$. The expected utility is therefore given by:

$$EU = \int_{\underline{\delta}}^{\overline{\delta}} \left[A + \delta - T(A+\delta) - \eta\left(\frac{A}{v}\right)\right] D(\delta)\, d\delta = A - \int_{\underline{\delta}}^{\overline{\delta}} T(A+\delta) D(\delta)\, d\delta - \eta\left(\frac{A}{v}\right). \tag{17}$$

We call $\hat{T}'(A) = \int_{\underline{\delta}}^{\overline{\delta}} [d\hat{T}(A+\delta)/dA] D(\delta) d\delta$ the expected marginal tax rate. The first-order condition of the utility maximisation programme is:

$$v'\left(\frac{A}{v}\right) = v\left[1 - \hat{T}'(A)\right]. \tag{18}$$

We see that the optimum A depends on the distribution of δ, through the expected net-of-tax wage rate. When they make their choices, individuals do not consider the actual tax schedule, but the expected smooth one.

5.2 Parameterization

The overall concavity of the US tax system has decreased quite substantially during the last four decades (see Figure 3). To illustrate the sensitivity of the marginal deadweight loss, marginal tax revenue and marginal deadweight loss per dollar, we provide numerical computations for the US economy in 1979, 1993 and 2006. These computations should be regarded as illustrative.

We take into account the federal income tax, the state income tax, the earned income tax credit, the payroll tax, the state sales tax and the local sales tax. We use the Californian tax schedule to compute the state taxes. The disutility of effort η is given by:

$$\eta\left(\frac{A}{v}\right) = \frac{v}{1+1/\beta}\left(\frac{A}{v}\right)^{1+1/\beta}. \tag{19}$$

v can be interpreted as a wage parameter and β as the elasticity of taxable income. The distribution of v is obtained by inversion, using TAXSIM and the CPS labor extracts

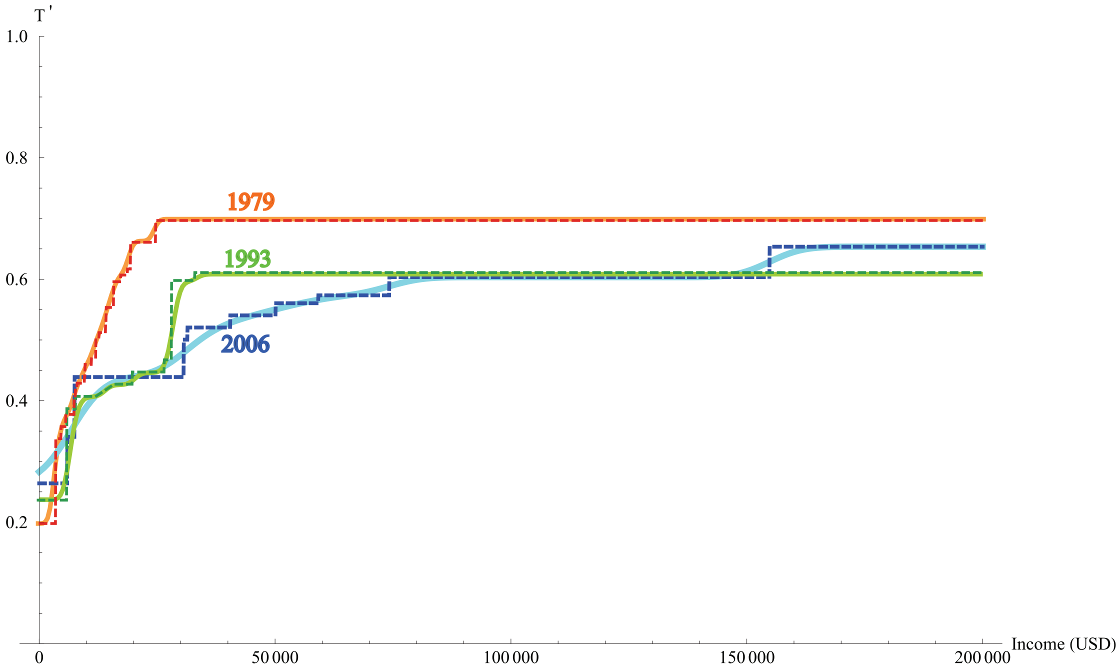


Figure 3: Statutory and expected marginal tax rates (plain and dotted lines respectively, benchmark calibration of σ)

restricted to single males for 1979, 1993 and 2006.

The standard deviation σ of the error term δ plays an important part in the computations. There is however no obvious way to calibrate it. We assume that it is normally distributed and that individuals face a positive shock larger than 10% of average gross income with a probability of 10% (and the same for a negative shock). We obtain $\sigma = 859$ for 1979, $\sigma = 1,227$ for 1993 and $\sigma = 2,787$ for 2006. We below examine the sensitivity of our numerical results with respect to this parameter.

We use a Hermite interpolation of order two in order to approximate the expected tax schedule. This means that the interpolation matches the unknown function both in observed values, and the observed values of its first two derivatives. This procedure is satisfying as our results depend one the slope and on the curvature of the expected tax schedule. Figure 3 shows the statutory piecewise linear schedule and the corresponding smooth schedule in 1979, 1993 and 2006.

5.3 Numerical Results

In the tables below, we show the bias in various measures of the deadweight loss. One measure is the deadweight loss per marginal tax dollar. It is often used and easy to understand as long as we are on the left hand side of the Laffer curve. However, when we approach the top of the Laffer curve, the MTR will go to zero and the marginal deadweight loss per marginal tax dollar to infinity. Then, for further increases in the taxable income elasticity, the marginal tax revenue will be negative and the marginal deadweight loss per marginal tax dollar will be negative. The intuition for such a negative number is less straightforward than when the marginal tax revenue is positive. For this reason, we do not report the marginal deadweight loss per marginal tax dollar for values of the taxable income elasticity where we are on the right-hand side of the Laffer curve.

Tables 1, 2 and 3 summarize the results, for our benchmark calibration of σ and elasticities in the range 0.2–1.0. The marginal deadweight loss and tax revenue that we report correspond to a 10% increase in t. For example, we see that the linearization procedure leads to an overestimation of the marginal deadweight loss by 15.4% in 1979, 9.7% in 1993 and 6.3% in 2006, for an elasticity $\beta = 0.4$. These biases are much larger for an elasticity of 1.0 – equal to 38.2%, 19.1% and 13.7% respectively. Regarding the marginal deadweight loss per marginal tax dollar, it clearly appears that the biases in the deadweight loss and tax revenue do not alleviate, but reinforce each other. During the last forty years, and until the financial crisis of 2007, there has been in many countries a tendency to reduce marginal tax rates faced by high income earners. This has made the tax systems less curved overall. In these cases, the overestimation of the marginal deadweight loss and the underestimation of the marginal tax revenue are relatively small. However, marginal tax rates faced by high income earners have been increased in the last few years and further increases are vividly discussed. In this context, it is critical to correctly account for the curvature of the tax function.

In Tables 4 and 5, we examine the sensitivity of our calibration with respect to changes in the standard deviation σ of the error term δ, for an elasticity β of 0.4 and 0.8 respectively. We consider two deviations to the left and two deviations to the right, for year 2006. The results of our benchmark calibration are quite robust to relatively large variations in σ. For example, for $\beta = 0.4$, a multiplication of σ by 11, from \$500 to \$5,500, is associated with a 5%-change in the marginal deadweight loss and a 0.26%-change in the marginal tax revenue.

Table 1: Benchmark Case, Year 1979

β	0.2	0.4	0.6	0.8	1.0
DW	\$42	\$77	\$100	\$118	\$122
DW linear	\$45	\$88	\$124	\$154	\$169
Bias	*+7.8%*	*+15.4%*	*+23.6%*	*+30.1%*	*+38.2%*
TR	\$224	\$162	\$80	\$22	–
TR linear	\$221	\$150	\$60	–	–
Bias	*-1.4%*	*-7.3%*	*-29.7%*	–	–
DW per \$	\$0.19	\$0.47	\$1.26	\$5.50	–
DW per \$ linear	\$0.20	\$0.59	\$2.21	–	–
Bias	*+9.3%*	*+24.5%*	*+75.9%*	–	–

Table 2: Benchmark Case, Year 1993

β	0.2	0.4	0.6	0.8	1.0
DW	\$27	\$47	\$61	\$68	\$69
DW linear	\$28	\$51	\$69	\$79	\$82
Bias	*+5.2%*	*+9.7%*	*+13.9%*	*+16.0%*	*+19.1%*
TR	\$74	\$36	\$6	–	–
TR linear	\$73	\$32	–	–	–
Bias	*-1.9%*	*-12.5%*	–	–	–
DW per \$	\$0.36	\$1.28	\$10.3	–	–
DW per \$ linear	\$0.39	\$1.61	–	–	–
Bias	*+7.2%*	*+25.3%*	–	–	–

Table 3: Benchmark Case, 2006

β	0.2	0.4	0.6	0.8	1.0
DW	\$131	\$237	\$311	\$363	\$386
DW linear	\$136	\$252	\$340	\$404	\$439
Bias	*+3.6%*	*+6.3%*	*+9.1%*	*+11.5%*	*+13.7%*
TR	\$554	\$376	\$210	\$75	–
TR linear	\$549	\$361	\$182	\$33	–
Bias	*-0.8%*	*-4.0%*	*-13.5%*	*-55.6%*	–
DW per \$	\$0.24	\$0.63	\$1.48	\$4.84	–
DW per \$ linear	\$0.25	\$0.70	\$1.87	\$12.2	–
Bias	*+4.5%*	*+10.7%*	*+26.1%*	*+151.1%*	–

Table 4: Year 2006, Variation in σ for $\beta = 0.4$

σ	\$500	\$1,500	\$2,787 Benchmark	\$4,500	\$5,500
DW	\$247	\$242	\$237	\$235	\$235
DW linear	\$256	\$253	\$252	\$252	\$254
Bias	*+3.4%*	*+4.5%*	*+6.3%*	*+7.6%*	*+8.0%*
TR	\$373	\$378	\$376	\$372	\$372
TR linear	\$364	\$367	\$361	\$355	\$354
Bias	*-2.2%*	*-2.9%*	*-4.0%*	*-4.8%*	*-5.1%*
DW per \$	0.65	0.64	0.63	0.63	0.63
DW per \$ linear	0.69	0.69	0.70	0.71	0.72
Bias	*+5.8%*	*+7.7%*	*+10.7%*	*+13.0%*	*+13.8%*

Table 5: Year 2006, Variation in σ, for $\beta = 0.8$

σ	\$500	\$1,500	\$2,787 Benchmark	\$4,500	\$5,500
DW	\$362	\$357	\$363	\$354	\$332
DW linear	\$401	\$398	\$404	\$406	\$385
Bias	*+10.9%*	*+11.4%*	*+11.5%*	*+14.9%*	*+15.9%*
TR	\$73	\$72	\$75	\$76	\$57
TR linear	\$33	\$31	\$33	\$23	\$4
Bias	*-54.3%*	*-56.9%*	*-55.6.0%*	*-69.7%*	*-93.4%*
DW per \$	4.96	4.98	4.84	4.67	5.87
DW per \$ linear	12.03	12.87	12.2	17.71	103.9
Bias	*+142.5%*	*+158.4%*	*+10.7%*	*+279.6%*	*+1670.3%*

6 Conclusion

Actual tax systems are usually such that the marginal tax changes with the income level, implying that the budget constraints that individuals face are nonlinear. It is of interest to calculate the marginal deadweight loss of changes in a nonlinear income tax. A nonlinear income tax can be varied in many different ways. Break points can be changed, the intercept can be changed and the slope can be changed. Moreover, the slope can be changed in different ways. We do not cover all these different possibilities to vary a nonlinear tax. We focus on a particular kind of change in the slope, namely a change in the slope such that the marginal tax changes with the same number of percentage points at all income levels. Such a change can represent, for example, a change in the pay roll tax, the value added tax or a proportional state income tax. A common procedure to calculate the marginal deadweight loss of a change as described above has been to linearize the budget constraint at some point and then calculate the marginal deadweight loss for a variation in the linearized budget constraint. As shown in the article, such a procedure does not give the correct value of the marginal deadweight loss.

In this article, we derive the correct way to calculate the marginal deadweight loss when the budget constraint is smooth and convex. It is well known that the size of the deadweight loss depends on the curvature of the indifference curves, with more curved indifference curves yielding smaller substitution effects and lower marginal deadweight losses. We show that the curvature of the budget constraint is equally important for the size of the marginal deadweight loss. In fact, the curvature of the budget constraint enters the expression for the marginal deadweight loss in exactly the same way as the curvature of the indifference curve. We then numerically show that the bias introduced by the linearization procedure is often quite large, for reasonable parameter values.

Because of the potentially misleading policy implications of the linearization procedure, we believe that the curvature of the tax system should be fully accounted for when measuring the deadweight loss.

References

2010. "Optimal Redistributive Taxation with Both Extensive and Intensive Responses." IZA Discussion Papers 4837, Institute for the Study of Labor (IZA). URL `http://EconPapers.repec.org/RePEc:iza:izadps:dp4837`.

Auerbach, A. and J. Hines. 2002. *Handbook of Public Economics*, vol. 3, chap. Taxation and economic efficiency. Elsevier, 1347 – 1421.

Blomquist, S. 1983. "The effect of income taxation on the labor supply of married men in Sweden." *Journal of Public Economics* 22:169–197.

Blomquist, S. and W. Newey. 2002. "Nonparametric Estimation with Nonlinear Budget Sets." *Econometrica* 70:2455–2480.

Burtless, G. and J. A. Hausman. 1978. "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment." *Journal of Political Economics* 86:1103–1130.

Chetty, R. 2012. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica* 80:969–1018.

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

Dupuit, J. 1844. "De la mesure de l'utilite des travaux publics." 2:332–375.

Eissa, N., H. Kleven, and C. Kreiner. 2008. "Evaluation of four tax reforms in the United States: Labor Supply and welfare effects for single mothers." *Journal of Public Economics* 92:795–816.

Feldstein, M. 1995. "Behavioral Responses to Tax Rates: Evidence from TRA86." *American Economic Review* 85:170–174.

———. 1999. "Tax Avoidance and the Deadweight Loss of the Income Tax." *Review of Economics and Statistics* 81:674–680.

Gelber, A. and J. W. Mitchell. 2012. "Taxes and Time Allocation: Evidence from Single Women and Men." *Review of Economic Studies* 79:863–897.

Gruber, J. and E. Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84:1–32.

Harberger, A. 1962. "The incidence of the corporation income tax." *Journal of Political Economy* 70:215–240.

———. 1964. "The Measurement of Waste." *American Economics Review* 54:58–76.

Hausman, J. A. 1979. "The econometrics of labor supply on convex budget sets." *Economics Letters* 3:171–174.

Hausman, Jerry A. 1985. "The Econometrics of Nonlinear Budget Sets." *Econometrica* 53 (6):pp. 1255–1282.

Kleven, H. and C. Kreiner. 2006. "The marginal cost of public funds: Hours of work versus labor force participation." *Journal of Public Economics* 90:1955 – 1973.

Kopczuk, W. 2005. "Tax bases, tax rates and the elasticity of reported income." *Journal of Public Economics* 89:2093–2119.

Saez, E. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68:205–229.

Saez, E., J. Slemrod, and S. Giertz. 2012. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50. Forthcoming in the Journal of Economic Literature.