Marginal Deadweight Loss when
the Income Tax is Nonlinear

Sören Blomquist and Laurent Simula

Uppsala University and Uppsala Center for Fiscal Studies[*][†]

March 8, 2012

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. This is quite surprising since most income tax systems are nonlinear, generating nonlinear budget constraints. Instead of developing the proper procedure to calculate the marginal deadweight loss for variations in nonlinear income taxes a common procedure has been to linearize the nonlinear budget constraint and apply methods that are correct for variations in a linear income tax. Such a procedure leads to incorrect results. The main purpose of this paper is to show how to correctly calculate the marginal deadweight loss when the income tax is nonlinear. A second purpose is to evaluate the bias in results that obtains when a linearization procedure is used. We perform calculations for the US tax system in 1979, 1994 and 2006. We find that the linearization procedure significantly overestimates the marginal deadweight loss and underestimates the marginal tax revenue. The magnitude of the errors has decreased over the last three decades, reflecting the reduction in the curvature of the US tax schedule.

Keywords: Deadweight Loss, Taxable Income, Nonlinear Budget Constraint

JEL classification: H21, H24, H31, D61

I.   Introduction

The study of the deadweight loss 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 fifties and sixties (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, 2010; Saez, Slemrod, and Giertz, 2009). These taxable income functions show how taxable income varies as the slope of a linear budget constraint of individuals is changed at the margin.

Most theoretical work on how to calculate the marginal deadweight loss has been done for linear taxes and hence for variations in linear budget constraints. This is quite surprising because most income tax systems are nonlinear, generating nonlinear budget constraints. Instead of developing the proper procedure to calculate the marginal deadweight loss for variations in nonlinear income taxes, one has linearized the nonlinear budget constraint and applied the procedure that is correct for variations in a linear income tax. Empirical works usually linearize budget constraints and proceed as if the budget constraints were linear. As we will show, this leads to incorrect results. The main purpose of our article is to show how to correctly calculate the marginal deadweight loss when the income tax is nonlinear. A second purpose is to evaluate the bias in results that obtains when the budget constraint is linearized and the linearized budget constraint is used to calculate the marginal deadweight loss. For tax systems where the marginal income tax increases with the taxable income, this linearization procedure may often lead to an overestimate of the marginal deadweight loss.

Actual tax systems are usually piecewise linear and, in the end, we describe how to calculate the marginal deadweight loss for such tax systems. However, in order to get simple and clean results, we start our analysis by considering smooth budget constraints. It should be noted that the average, or aggregate, behavior for a population does not depend on whether the tax system and budget constraints are kinked or smooth. It is

the general shape of the tax system and budget constraints that determine the average behavior. To further simplify the analysis, we consider tax systems that generate convex budget sets. However, in our numerical computations, we fully account for nonconvexities. Historically, much focus has been on how the income tax distorts labor supply. Since the more recent literature has the focus on taxable income, we state our results in terms of this concept. Of course, it is easy to modify our results to some other application.

When doing our theoretical analysis, as well as numerical calculations, we will do this for a change in the tax schedule such that individuals' budget constraints rotate downwards and such that the marginal tax rate increases with the same number of percentage points at all income levels. This is the kind of experiment considered by Feldstein is his seminal (1999) article. Such a change can, for example, be interpreted as an increase in the payroll tax. Our analysis and numerical calculations are done for the marginal deadweight loss, the marginal tax revenue and the marginal deadweight loss per marginal tax dollar. If the tax system is such that the marginal tax increases with income, the linearization procedure overestimates the change in marginal deadweight loss and underestimates the change in tax revenue. When the overestimated marginal deadweight loss is divided by the underestimated marginal tax revenue to get the marginal deadweight loss per dollar, the two mistakes are magnified so that the bias in marginal deadweight loss per dollar in some cases becomes very large.

As will be shown below, the expression for the marginal deadweight loss when a linearization procedure is used looks formally the same as the correct expression. The crucial difference is that if the linearization procedure is used one misses the fact that the change in taxable income depends on the curvature of the budget constraint. In fact, the curvature of the budget constraint and the curvature of the indifference curve are of equal importance for how large the change in taxable income will be. If the budget constraint is linearized one sets the curvature to zero and overestimates the change in taxable income.

Actual tax systems are usually piecewise linear. Unless the taxable income elasticity is very small we would expect to see a fair amount of bunching at the kink points. However, such bunching is usually not observed. Early studies in the nonlinear budget set literature (e.g., Burtless and Hausman (1978), Hausman (1979), Blomquist (1983) and Hausman (1985)) and more recently Chetty (2009) have emphasized that, due to optimization errors, usually there will be a difference between desired taxable income (hours of work) and realized taxable income. These optimization errors imply that even if there would be bunching at kink points of desired taxable income, we should not expect to see much bunching of actual taxable income.

To evaluate how misleading results will be if one uses linearized budget constraints to calculate the marginal deadweight loss, we perform calculations on the US tax system for the years 1979, 1994 and 2006. The overall curvature of the tax schedule differs quite much between these years. The bias in the linearization procedure depends both on the general curvature of the tax system and the curvature of the indifference curves, i.e., on the taxable income elasticity. Our calculations range from a low value of the bias in the marginal deadweight loss per tax dollar of 4.1% for the 2006 tax system and an elasticity of 0.2 to a bias of 132% for the 1979 tax system and an elasticity of 0.8. [1]

Feldstein (1999) provides computations of the marginal deadweight loss of the US tax system in 1994. He computes a marginal deadweight loss per tax dollar of $2.06. Using the same set of taxes, the same year and the same elasticity of taxable income, we are able to replicate this number closely; we obtain $2.16, when using the linearization procedure. However, when properly taking account of the curvature of the budget constraint, we calculate the marginal deadweight loss per tax dollar to be $1.35. That is, the linearization procedure overestimates the marginal deadweight loss per tax dollar by 61%.

It is simplest to illustrate the basic ideas under the assumption that the budget constraints are smooth. In Section II, we therefore use smooth budget constraints to introduce the main idea and show how the marginal deadweight loss should be correctly calculated. We also give expressions for how large the bias using a linearization procedure can be. In reality, tax systems normally create piecewise linear budget constraints. In Section III, we show how the calculations are modified if budget constraints are piecewise linear. In Section IV, we present calculations of the marginal deadweight loss for the US tax system for three different years. Section V concludes.

II.   Smooth Income Tax

Actual tax systems are usually piecewise linear and, in the next section, we describe how to calculate the marginal deadweight loss for such tax systems. However, in order to get simple and clean results, we start our analysis by considering smooth budget constraints. One reason is that it is basically the general shape of the tax system and the budget constraints that determine the average behavior.[2]

---

[1] In fact, we obtain a bias that is even more severe if the elasticity is 1.0. However, because the linearization procedure in that case incorrectly indicates that we are on the wrong side of the Laffer curve the percentage figure is not comparable to the other percentage figures.

[2] This should be qualified. A smooth tax schedule is a good approximation of a piecewise linear tax schedule provided the distribution of the kink points is regular enough.

A. 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 matter 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 $t \geqq 0$. Note that in this case $\partial^2 T(A) / \partial A^2$ reduces to $g''(A)$. 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. With the parametrization we use, 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.[3] A second reason is, of course, that real tax systems are, as a first approximation, of a form as the one described by $g(A) + tA$.

---

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

B. The True Marginal Deadweight Loss

Consider the utility maximization problem:

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

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 denote the solution to problem (P1) as $C(t, B, v), A(t, B, v)$. The form of these two functions depends on the functional forms of U and g. Sticking $C(t, B, v), A(t, B, v)$ 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 t. We also examine the marginal deadweight loss per marginal tax dollar, as this measure is often used in the literature. We first derive the correct expressions and then – in the next subsection – describe how they usually are calculated. For this purpose, we define the expenditure function as:

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

This problem also defines the compensated demand and supply functions, $C^h\left(t, v, \overline{u}\right)$ and $A^h\left(t, v, \overline{u}\right)$ 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 $U\left(C, A, v\right)$ *and on the functional form of* $g\left(A\right)$. 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.

Let us define the compensated tax revenue function as:

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

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 the difference between compensated changes in expenditure and collected taxes, i.e.,

$$MDW := \frac{dE\left(t,v,\overline{u}\right)}{dt} - \frac{dR(A^h\left(t,v,\overline{u}\right))}{dt}$$

$$= A^h - g'(A^h)\frac{dA^h}{dt} - A^h - t\frac{dA^h}{dt} = -(g'\left(A^h\right) + t)\frac{dA^h}{dt}, \tag{3}$$

where we used the envelope theorem to obtain $dE\left(t,v,\overline{u}\right)/dt = A^h$. Expression (3) is the correct expression for the marginal deadweight loss. The ratio (3)/(2) corresponds to the marginal deadweight loss per marginal tax dollar.

C. A Commonly Used Linearization Procedure

We next describe a commonly used procedure that, in general, overestimates the marginal deadweight loss.[4] Let us consider particular values v^*, t^* and B^* and the solution to (P1), $C^* = C\left(t^*, v^*, B^*\right), A^* = A\left(t^*, v^*, B^*\right)$. 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 $M = C^* - A^* + p_A A^*$. 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

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

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:

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

---

[4]Feldstein (1999) is an example of an empirical study using a linearization procedure and (Dahlby, 1998) is an example of a theoretical study using a linearization procedure.

Figure 1: Nonlinear and linearized programmes
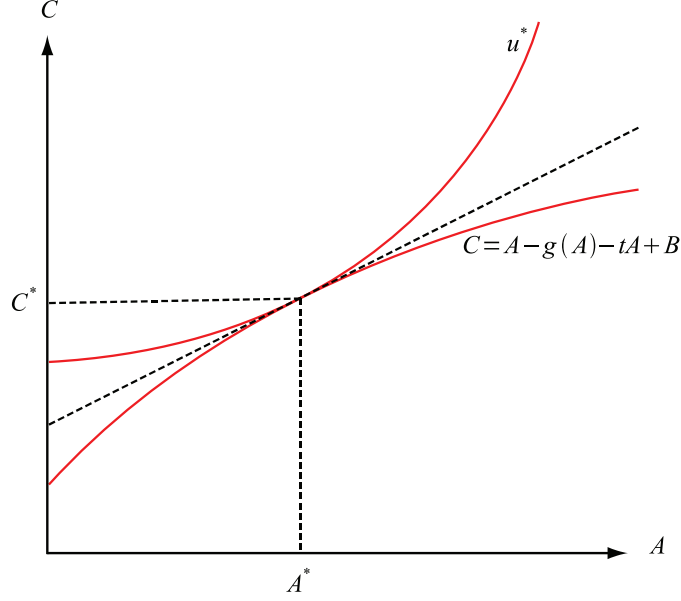
We obtain the marginal deadweight loss as:

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

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

$$A\left(t,B\right) \equiv A^h\left(U\left(C\left(t,B\right),A\left(t,B\right)\right)\right)$$
$$\equiv A_L\left(p_A\left(C\left(t,B\right),A\left(t,B\right)\right),M\left(C\left(t,B\right),A\left(t,B\right)\right)\right) \equiv A_L^h\left(U\left(C\left(t,B\right),A\left(t,B\right)\right)\right). \quad (6)$$

D. Comparative Statics for Taxable Income

Expressions (3) and (5) obtained above 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.

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 $T\left(A\right) = tA + pA + \pi A^2$, where we can interpret tA as the state tax and $pA + \pi A^2$ as the federal tax. This yields a budget constraint $C = A - \left(p + t\right)A - \pi A^2 + B$, where B is lump-sum income. Substituting the budget constraint into the utility function, we obtain $U = A - \left(p + t\right)A - \pi A^2 + B - \alpha A - \gamma A^2$. Maximizing with respect to A, we get $dU/dA = 1 - \left(p + t\right) - 2\pi A - \alpha - 2\gamma A$. We see that a necessary condition for a nonnegative A is $1 - \left(p + t\right) - \alpha \geqq 0$. We find that $d^2U/dA^2 = -2\left(\pi + \gamma\right) < 0$ for $\pi + \gamma > 0$. Setting $dU/dA = 0$ and solving for A, we obtain

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

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

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

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 $2\left(\pi + \gamma\right)$ 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 2, 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 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

10

Figure 2: Deadweight loss when the budget constraint is nonlinear (left panel) and linearized (right panel)
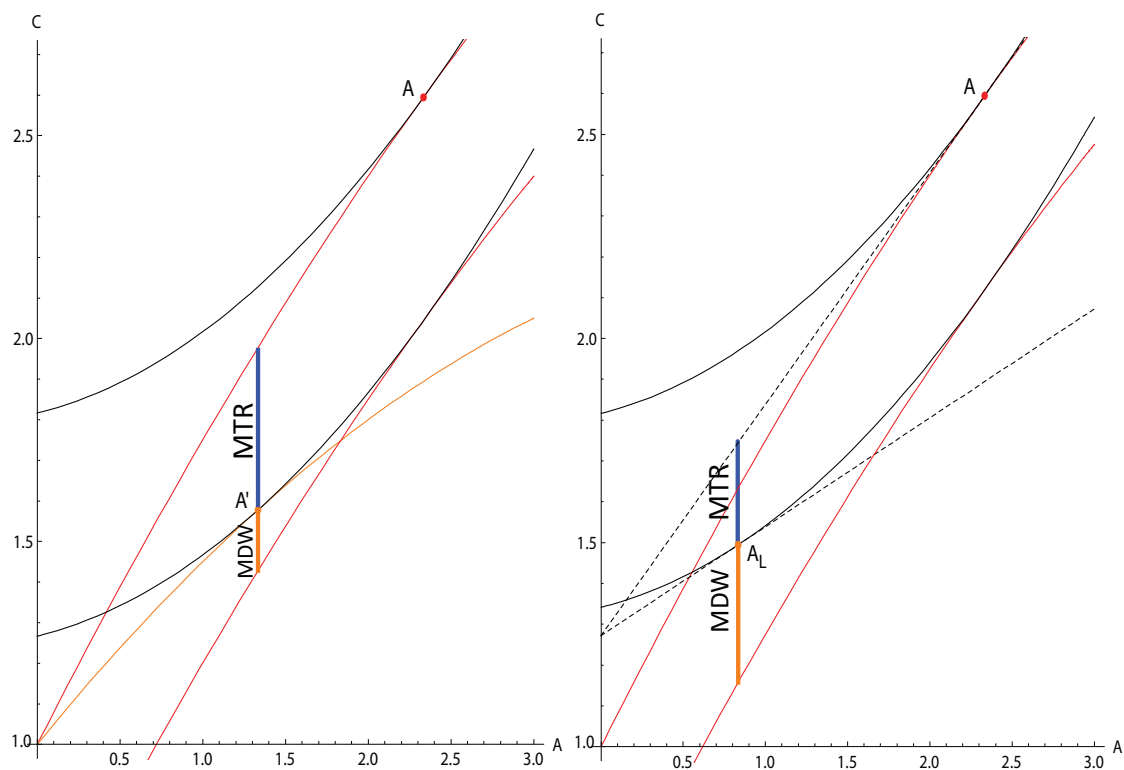
employs a variation in the linearized budget constraint. The nonlinear budget constraint through A is linearized around this point. The increase in t induces a rotation of the linearized budget constraint around the intercept. The bundle A_L is now chosen instead of A. We see that the deadweight loss, 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.

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 $U(C, A, v) \geqq \overline{u}$ is binding at the op-

timum 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.* 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 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:

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

E. 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

12

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.*

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'' also 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:

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

In tables presented in section IV, we show the bias in various measures of deadweight loss if one uses linearized budget constraints. One measure is the deadweight loss per marginal tax dollar. This is an often used measure and it is easy to understand as long as we are on the left hand side of the Laffer curve. In the tables, we will see how this measure changes as we increase the elasticity of taxable income. As 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 MTR 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 MTR 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. This does not mean that it is less serious to linearize the budget constraint. It is really more serious. If one uses a linearized budget constraint, this procedure will indicate the top of the Laffer curve ($MTR = 0$) at a lower value of the taxable income elasticity than what is correct.

For values of the taxable income elasticity such that we are on the left-hand side of

the Laffer curve, we present the bias in the marginal deadweight loss per marginal tax dollar implied by the linearization procedure computed as:

$$\frac{MDW_L/MTR_L}{MDW/MTR} = \frac{MDW_L/MTR_L}{MDW/(A-MDW)} = \frac{(1+a)\,A-MDW_L}{MTR_L} = 1 + \frac{a\,A}{MTR_L}. \quad (16)$$

Once again, we see that a, the relative curvature between the indifference curve and the budget constraint, is a key determinant of how serious the bias is.

III. Piecewise Linear Income Tax

In reality, tax systems normally create piecewise linear budget constraints. We now investigate how the results obtained above are modified if budget constraints are piecewise linear. More specifically, the tax system that we consider is of the same form as above, but the federal tax is piecewise linear. To illustrate the mechanisms at work, it is sufficient to consider a tax system generating a budget constraint with two linear segments and one kink point. The results easily generalize to a tax system with many kinks.[5]

Let the federal tax system be characterized by the marginal tax rate τ_1 for taxable income up to the break point A_1 and the marginal tax τ_2 for incomes above the break point. Let there also be a state income tax of t. The budget constraint generated by this tax system is shown in Figure 3.

The slope on the first segment is given by $\theta_1 = 1 - \tau_1 - t$ and on the second segment by $\theta_2 = 1 - \tau_2 - t$. The intercept for the first segment, R_1, is lump-sum income. The virtual income for the second linear segment is given by $R_2 = R_1 + (\theta_1 - \theta_2)\,A_1 = R_1 + (\tau_2 - \tau_1)\,A_1$. Hence, the latter does not depend on t and does not change when t is varied.

A. Individual Behavior

To make the problem interesting, we need some individuals locating in the interior of the segments and some at the kink point. Hence, we now re-introduce the heterogeneity parameter v explicitly and write the utility function $U(C, A, v)$, where v is a preference

---

[5]Dahlby (1998) studies how the marginal cost of public funds should be calculated when the income tax schedule is piecewise linear. Like us, he considers the case with heterogenous individuals. However, in his analysis, he does not realize and does not take into account that some individuals will have their desired hours of work at kink points and that when there is a marginal change in a tax parameter these individuals will not change their hours of work.

Figure 3: Piecewise Linear Budget Constraints
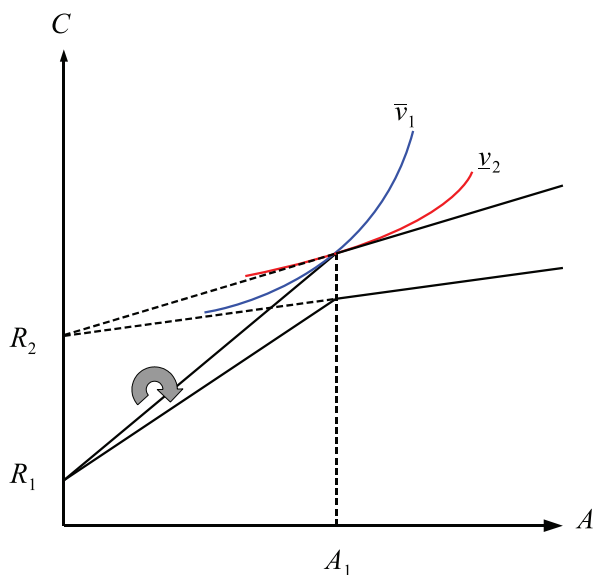
parameter with pdf $\phi(v)$ over $(\underline{v}, \overline{v})$. If we had a pure labor supply model, it would be natural to write the utility function as $U(C, A/v)$ and interpret v as the wage rate. However, since in the taxable income literature it is assumed that there are other margins than hours of work, we prefer to write it in the more general form $U(C, A, v)$. For one interval of v, we would have solutions on the first segment, for another interval at the kink point and, for a third interval, on the second segment.

A first step is to find out how the budget constraint changes as the tax parameter t increases. We know that R_1 and R_2 do not change. The slopes of the first and second segments decrease. The kink point is still at A_1. However, its C-coordinate decreases by $dt \times A_1$, the amount of the extra tax paid.

For a person located at the kink point before and after the change in t, we have $dA^h/dt = 0$. Therefore, there is no marginal deadweight loss from the increase in t for this person. The increase in taxes paid by a person located at the kink is just like a lump-sum tax.

For a person with a tangency on one of the linear segments, the variation in the budget constraint is just like a variation in a linear budget constraint. For such a person, one can therefore apply the taxable income function that is generated by a linear budget constraint and the marginal deadweight loss for an individual with parameter v would

be given by

$$- (\tau_i + t) \frac{dA_L^h (\tau_i + t, v, \overline{u})}{dt}, \tag{17}$$

with $i = 1, 2$, where we should remember that $\overline{u}$ is a function of v.

The expressions for the marginal tax revenue are quite straightforward. For a person with desired taxable income located in the interior of the segment i, the marginal tax revenue is given by $MTR = A + (\tau_i + t) \ dA/dt$ whereas we obtain $MTR = A$ for a person with desired taxable income at the kink point.

B.   Marginal Deadweight Loss for the Population

If we want to find the aggregate marginal deadweight loss, we can integrate over v. For simplicity, we assume that v enters the utility function in such a way that A_L^h is strictly increasing in v. We also assume that $0 < A_L (\tau_1 + t, \underline{v}) < A_1 < A_L (\tau_2 + t, \overline{v})$. Hence, no one chooses the zero solution and there are individuals choosing a bundle on the first segment, some others at the kink and some others on the second segment. Let $\overline{v}_1$ be defined by $A_L (\tau_1 + t, \overline{v}_1) = A_1$ and $\underline{v}_2$ by $A_L (\tau_2 + t, \underline{v}_2) = A_1$ as shown in Figure 3. Define the subsets $S_1 = (\underline{v}, \overline{v}_1)$ and $S_2 = (\underline{v}_2, \overline{v})$. Likewise define the set $K_1 = (\overline{v}_1, \underline{v}_2)$. Then individuals with $v \in S_1$ will have a solution on the first segment, individuals with $v \in S_2$ on the second segment and persons with $v \in K_1$ a solution at the kink point.

The aggregate (non-marginal) deadweight loss is an expression that can be written as

$$\int_{S_1} \delta_1 (t, v) \phi (v) \, dv + \int_{K_1} \delta_2 (t, v) \phi (v) \, dv + \int_{S_2} \delta_3 (t, v) \phi (v) \, dv, \tag{18}$$

where δ generically represents the (non-marginal) deadweight loss for person v. The aggregate marginal deadweight loss is the derivative of this expression with respect to t, which by Leibnitz's rule is equal to

$$\int_{S_1} \delta'_{1t} (t, v) \phi (v) \, dv + \overline{v}'_1 (t) \delta_1 (t, \overline{v}_1 (t)) \phi (\overline{v}_1 (t))$$

$$+ \int_{K_1} \delta'_{2t} (t, v) \phi (v) \, dv + \underline{v}'_2 (t) \delta_2 (t, \underline{v}_2 (t)) \phi (\underline{v}_2 (t)) - \overline{v}'_1 (t) \delta_2 (t, \overline{v}_1 (t)) \phi (\overline{v}_1 (t))$$

$$+ \int_{S_2} \delta'_{3t} (t, v) \phi (v) \, dv - \underline{v}'_2 (t) \delta_3 (t, \underline{v}_2 (t)) \phi (\underline{v}_2 (t)). \tag{19}$$

To compute the (non-marginal) deadweight loss for person $\overline{v}_1$ or $\underline{v}_2$, we need to consider his taxable income and his highest feasible indifference curve. So, it is not the slope of the budget constraint that matters here. The slope of the budget constraint matters when

16

evaluating his marginal deadweight loss. We thus have $\delta_1(t, \overline{v}_1(t)) = \delta_2(t, \overline{v}_1(t))$ and $\delta_2(t, \underline{v}_2(t)) = \delta_3(t, \underline{v}_2(t))$, which implies that the terms in (19) showing movements in and out of the kink point sum to zero. Consequently, the aggregate marginal deadweight loss is

$$MDW = -\sum_{i=1}^{2} \int_{S_i} (\tau_i + t) \frac{dA_L^h}{dt} \phi(v)\, dv + 0 \times \int_{K_1} \phi(v)\, dv. \tag{20}$$

The aggregate marginal tax revenue is given by:

$$MTR = \sum_{i=1}^{2} \int_{S_i} \left[A + (\tau_i + t) \frac{dA_L^h}{dt} \right] \phi(v)\, dv + \int_{K_1} A\, \phi(v)\, dv. \tag{21}$$

The contribution to the aggregate marginal deadweight loss from those at a kink point is zero. The difference between the smooth case and the piecewise linear case is that, in the former, the actual marginal deadweight loss is lower than that indicated by the "linear" taxable income function for any v and the corresponding value of A. In the piecewise linear case, the difference in the two measures is concentrated to the kink. If there were several kinks, the difference would also be concentrated to the kinks.

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 the average or total marginal deadweight loss as we just did. 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. 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.

IV.   How Serious is It to Linearize? Numerical Computations

We now provide computations of the marginal deadweight loss for the US tax systems in 1979, 1994 and 2006. We choose 1979 to emphasize the important part played by the curvature of the tax system, as the US taxes were significantly more progressive three

17

decades ago. We also consider 1994 as this is the year considered by Feldstein (1999).

The federal income tax is piecewise linear and the resulting budget constraint is of the form analyzed in the previous section. According to the analysis in that section there should be individuals with desired taxable income at kink points. Before progressing with our computations, we have to reflect a bit on the fact that looking at actual data there is very little bunching at kinks of the US tax schedule.

A. Why Are There so Few Individuals Observed at Kink Points?

The analysis above shows the importance of kink points. Observed wage distributions, together with budget constraints generated by actual tax systems and (estimated) compensated elasticities, often imply that a significant number of individuals have their desired hours at a kink point. This seems to be in conflict with the observation that very few individuals locate at kink points.

Using microdata from US tax returns over the period 1960-97, Saez (2010) finds clear evidence of bunching around the first kink of the Earned Income Tax Credit among self-employed workers and, to a lesser extent, around the threshold of the first tax bracket where tax liability starts. He finds little evidence of bunching at other tax brackets. Other studies have found modest evidence of bunching, for elderly US workers who are both working and receiving social security benefits (Burtless and Moffitt (1984) and Friedberg (2000)), above the first eligibility threshold for the UK earned income tax credit (Blundell and Hoynes, 2004) or for the Australian Higher Education Contribution Scheme (Chapman and Leigh, 2009). Bastani and Selin (2011), studying the Swedish tax system, find very little bunching.

We see several possible explanations for the fact that few people are observed at kink points. A first explanation is that the compensated elasticity is much smaller than often assumed. A second explanation, which we believe is the most important one, is that there are optimization errors. As already mentioned above, Burtless and Hausman (1978), Hausman (1979), Blomquist (1983), Hausman (1985) and more recently Chetty (2009) have emphasized that there is usually a difference between desired taxable income and realized taxable income because of optimization errors or frictions. The latter imply that even if there would be bunching at kink points of desired taxable income, we should not expect to see much bunching of actual taxable income.

If the first interpretation is correct (lower compensated elasticity), marginal dead-weight losses are small and studies indicating large elasticities are incorrect. If the latter interpretation is correct, it means that the number of individuals that would be at a

kink point in the long run, when they have been able to fully adjust to the actual tax schedule, is underestimated in many studies. We will elaborate on this second point.

B.   How to Account for Optimization Errors

We see two different ways of accounting for optimization errors. In one, individuals are aware of the fact that there might be unforeseen shocks to their taxable income and take this into account when calculating their desired taxable income. In the other, individuals do not take the optimization errors into account in their decision making, because they lack sufficient information or face too high (re)optimization costs. In other words, an individual plans desired income and then faces randomness which introduces a gap between the planned and realized levels.

The second way of taking uncertainty into account was clearly described by Burtless and Hausman (1978, p. 1115): "Indexing individuals by i, we expect random differences to occur between observed hours supplied, H_i, and preferred hours of work, H_r. This random variation may be the result of measurement error, but a more important source of randomness arises because of unexpected variations in hours worked. Unexpected temporary layoffs, involuntary overtime, or short time due to cyclical downturns all provide potential reasons actual hours may diverge from 'normal' hours associated with a given job. These variations in hours are unanticipated by the individual and cause his actual hours H_i to differ from his preferred hours H_r".

Desired, or planned, taxable income is directly determined by the statutory piece-wise linear budget constraint. However, for various reasons desired taxable income can sometimes not be realized. For many objects of choice, the actual amount bought or sold would be equal to the desired quantity. For example, the actual number of dresses bought during a year is probably pretty close to the desired number of dresses. However, for taxable income there can be unexpected events (shocks) that make actual taxable income deviate from the desired or planned taxable income. Let us first give examples why actual taxable income might be lower than desired taxable income. The individual might plan for a given taxable income. However, because of unexpected sickness, layoff, new vacation plans because of a new love, etc., actual taxable income might be lower than the planned one. This effect of unexpected events would be larger the later in the tax year the event occurs. Taxable income might be higher than the income planned for because of vacation plans that are changed for some reason, better health than expected, assigned overtime, etc.

Let us write realized taxable income as $A^r = A^d + \epsilon$, where ϵ is the shock or opti-

mization error. Let us make the reasonable assumption that the shocks are independent of desired hours and the tax system. Then, $\partial A^r / \partial \zeta = \partial A^d / \partial \zeta$ for any tax parameter ζ of the tax system. This implies that for individuals with actual taxable income in the interior of a linear segment but with desired income at a kink point, neither desired nor actual taxable income will change as a response to a marginal change in a tax parameter and their marginal deadweight loss will be zero.

An implication of the above is that for some individuals with their taxable income in the interior of a linear segment, it is appropriate to calculate the marginal deadweight loss as for a linear budget constraint, while for some other individuals with their actual taxable income in the interior of a linear segment the marginal deadweight loss is zero. The reason is that individuals in the first group have their desired income on a linear segment while individuals in the second group have desired income at a kink point.

It is very important to recognize the part played by random shocks to taxable income (optimization errors) when computing the marginal deadweight loss and marginal tax revenue. Below we perform computations assuming that an individual plans desired income and then faces randomness.

C.   Calibration

We undertake our calculations of the marginal deadweight loss of the US tax system for three different years. 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 and restrict the analysis to single men with no children.[6] We use the Californian tax schedule to compute the state taxes. California is the state with the largest population and many other states have similar income tax schedules. The payroll tax and the sales taxes are linear (however, above a annually adjusted threshold, the payroll tax only consists of the Medicare tax of 2.9%). The payroll tax (FICA) is 15.3% in 2006 and 1994, and 12.26% for 1979. The Californian sales tax is 7.25%. Local taxes vary. In our computations, we assume that the local sales tax is 0.5%. Overall, the linear component of the tax system, denoted by t in the previous sections, is equal to 20.01% in 1979, and 23.05% in 1994 and 2006. Altogether, the budget constraint in the US exhibits nonconvexities in 1994 and 2006. They are sharp in 1994 and we therefore fully took them into account in our computations. By contrast, they are rather negligible in 2006. We thus convexified the budget constraint by taking its convex hull. We obtain budget constraints with 18, 20 and 13 kinks in 1979, 1994 and 2006 respectively.

---

[6]For 1979, the number of children is not available in the CPS dataset.We took the whole population of single men into account.

As in Saez (2010), we consider that individual preferences are described by a quasi-linear and isoelastic function of the form

$$U\left(C,A,v\right) = C - \frac{v}{1+1/\beta}\left(\frac{A}{v}\right)^{1+1/\beta}, \qquad (22)$$

where v can be interpreted as a wage parameter. The quasilinearity assumption implies that there is no income effect on taxable income and that the Marshallian and Hicksian taxable income functions are equal. The isoelastic assumption implies that the elasticity of taxable income is constant, equal to β. Given the preference specification, a v-person facing a linear budget constraint with slope θ has taxable income $A = v\theta^{\beta}$, where v is the annual wage rate.

The distribution of the heterogeneity parameter v is obtained from the CPS labor extracts 1979, 1994 and 2006. We assume that this parameter is distributed according to a lognormal distribution, that we calibrate so as to replicate the first two moments of the actual distribution. We proceed as follows to recover the underlying distribution of v from the distribution of realized incomes A^{r}. We know that, for a given person, desired income differs from realized income by an amount ϵ, i.e. $A^{r} = A^{d} + \epsilon$. Desired income is equal to $A^{d} = v\theta^{\beta}$ along a linear segment of the budget constraint with slope θ. The underlying wage rate v is therefore given by $v = \left(A^{r} - \epsilon\right)/\theta^{\beta}$. It is reasonable to assume that $\partial A^{r}/\partial\zeta = \partial A^{d}/\partial\zeta$ for any tax parameter ζ and $\mathbb{E}\left(\epsilon\right) = 0$. When calibrating v, we use the slope $\tilde{\theta}$ of the budget constraint faced by the individual with average realized income. Denoting $\sigma^{2} = \text{Var}\left(\epsilon\right)$, this implies that $\mathbb{E}\left(v\right) = \text{E}\left(A^{r}\right)/\tilde{\theta}^{\beta}$ and $\text{Var}\left(v\right) = \left(\text{Var}\left(A^{r}\right) - \sigma^{2}\right)/\tilde{\theta}^{2\beta}$. For example, for 2006 and $\beta = 0.2$, we get $\mathbb{E}\left(v\right) = 40,081$ and $\text{Var}\left(v\right) = 28,012^{2}$.

The standard deviation σ of the error term ϵ is a key element. For many occupations in the U.S. labor market, supplemental pay – overtime, bonuses, and shift differentials – is an important component of overall cash compensation. Overtime pay is especially important in production occupations and other blue-collar jobs; bonus pay is mostly a feature of high-wage managerial and sales occupations; and shift differentials play a prominent role in a particular set of occupations – healthcare practitioner and technical occupations. Regarding sick leave, private industry workers access to paid sick leave benefits varied by occupational group and ranged from 84% for management, professional, and related occupations to 42% for service workers.[7] For example, a worker who does not benefit from paid sick leave, taking the average number of sick days (14 days) and

---

[7]US Bureau of Labor Statistics, "On paid sick leave", *Program Perspective*, vol. 2, issue 2, March 2010

working 260 days/year (average number of working days in 2006) would incur of monetary loss of around 6% of his annual income. There is no obvious way to calibrate ϵ. We herein 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). Under these assumptions, we get $\sigma = 1,550/1.28 = 1,211$ for 1979, $\sigma = 2,095$ for 1994 and $\sigma = 2,787$ for 2006.

The size of the bias of the linearization procedure depends on the relative curvature of the budget constraint and the indifference curve. The latter is closely related to the taxable income elasticity; the larger the taxable income elasticity, the less curvature has the indifference curve. The curvature of the budget constraints is given by the tax systems of the three years, but how should we choose the taxable income elasticity? In the literature, there is a wide range of estimated labor supply and taxable income elasticities and there is no consensus about what the size of the elasticity really is. The only consensus that seems to exist is that the elasticity for secondary earners (usually women) is higher than that for primary earners (usually men).

Looking at the empirical literature that have used micro data to estimate labor supply elasticities, we find for men estimates of the compensated elasticity ranging from around 0.05 (Kosters, 1967) up to 1.22 (MaCurdy, 1983). These elasticities are for the intensive margin. Kimmel and Kniesner (1998) estimate both the intensive and extensive elasticity and find the total elasticity to be 1.25. For females the range for the intensive elasticity is from around 0.2 (Blundell, Duncan, and Meghir, 1998) up to around 2.0 (Blundell, Duncan, and Meghir, 1998) and extensive elasticities up to around 1.8 (Keane and Moffitt, 1998). Kimmel and Kniesner (1998) find a total elasticity for females of 3.05. Prominent macro economists argue that the labor elasticity is around 3.0 (Prescott, 2004; Rogerson and Wallenius, 2009). There is also a large literature on the taxable income elasticity. Blomquist and Selin (2010), using Swedish data estimates the taxable income elasticity for men to around 0.2 and around 1.0 to 1.5 for women. In general, the elasticities found in the taxable income literature are spanned by the labor supply elasticities given above. Given the wide variation in estimated elasticities, we calculate the bias that arises from linearizing for the following values of the elasticity: 0.2, 0.4, 0.6, 0.8 and 1.0.

D. Numerical Results

The linearization procedure ignores the issue of kink points. In order to compare it with our procedure, we allocate 50% of the individuals whose desired income would be at

Table 1: Results for 1979

β	0.2	0.4	0.6	0.8	1.0
% with desired A at a kink	7.1%	12.8%	18.1%	24.4%	28.9%
MDW correct	$22	$44	$66	$86	$106
MDW linear	$24	$50	$78	$109	$141
Bias	*+6.6%*	*+12.7%*	*+19.0%*	*+26.6%*	*+33.0%*
MTR correct	$102	$84	$67	$50	$33
MTR linear	$100	$79	$54	$27	-2
Bias	*-1.4%*	*-6.6%*	*-18.7%*	*-45.4%*	*–*
MDW per $	$0.22	$0.52	$0.98	$1.71	$3.18
MDW per $ linear	$0.23	$0.63	$1.43	$3.96	*–*
Bias	*+8.2%*	*+20.7%*	*+46.4%*	*+132.1%*	*–*

a kink to the linear segment to the left of the kink and 50% to the linear segment to the right, i.e. we assume that the distribution of random shocks is symmetric in the sense that half of the shocks are positive and half of them negative. In the linearization procedure, we treat those with desired hours at a kink point but actual hours on a linear segment as if desired taxable income was also on a linear segment.

Tables 1, 2 and 3 summarize the results. In 1979, the linearization procedure incorrectly indicate that we are on the downward sloping part of the Laffer curve when the elasticity is 1.0. In the tables, this is indicated by a hyphen. The linearization procedure overestimates the change in marginal deadweight loss and underestimates the change in tax revenue in a significant way. For an elasticity of 0.2, the marginal deadweight loss is overestimated by 6.6% and 4.7% in 1979 and 1994. This figure drops to 3.6% in 2006 because of the limited curvature of the budget constraint in that year. For larger values of the elasticity, the bias becomes more severe. For example, it is about 33% in 1979 for an elasticity of 1.0. We see that the magnitude of the error has decreased over the last three decades, reflecting the reduction in the curvature of the US tax schedule. The underestimation of the marginal tax revenue follows the same pattern. The two mistakes are magnified so that the bias in marginal deadweight loss per dollar in some cases becomes very large. Equal to 4.1% in 2006 for an elasticity of 0.2, it goes up to 132% in

Table 2: Results for 1994

β	0.2	0.4	0.6	0.8	1.0
% with desired A at a kink	3.8%	7.1%	9.6%	12.0%	13.2%
MDW correct	$29	$55	$78	$97	$116
MDW linear	$31	$60	$87	$113	$138
Bias	*+4.7%*	*+8.4%*	*+12.6%*	*+16.7%*	*+19.0%*
MTR correct	$168	$139	$114	$92	$71
MTR linear	$167	$134	$104	$76	$49
Bias	*-0.8%*	*-3.3%*	*-8.5%*	*-17.7%*	*-31.4%*
MDW per $	$0.17	$0.40	$0.68	$1.06	$1.63
MDW per $ linear	$0.18	$0.45	$0.84	$1.50	$2.83
Bias	*+5.6%*	*+12.1%*	*+23.1%*	*+41.8%*	*+73.4%*

Table 3: Results for 2006

β	0.2	0.4	0.6	0.8	1.0
% with desired A at a kink	3.4%	5.6%	6.6%	9.8%	12.0%
MDW correct	$33	$67	$100	$131	$162
MDW linear	$34	$70	$108	$145	$185
Bias	*+3.6%*	*+5.3%*	*+6.7%*	*+11.0%*	*+14.2%*
MTR correct	$246	$218	$190	$163	$137
MTR linear	$245	$215	$183	$149	$114
Bias	*-0.5%*	*-1.6%*	*-3.6%*	*-8.8%*	*-16.8%*
MDW per $	$0.13	$0.31	$0.53	$0.79	$1.18
MDW per $ linear	$0.14	$0.33	$0.59	$0.97	$1.62
Bias	*+4.1%*	*+7.0%*	*+10.6%*	*+21.6%*	*+37.2%*

1979 for an elasticity of 0.8.

Feldstein (1999) uses a linearization procedure to calculate the marginal deadweight loss for the income tax rates and rules of 1994 as described in TAXSIM and an elasticity of taxable income equal to 1.04. It can be of interest to see by how much this linearization biases his results. We have therefore replicated his calculations using both the correct and the linearization procedures.

There are three differences between his computations and the ones we provided above: Feldstein uses a triangle approximation of the deadweight loss while we used an exact computation; he also uses a different set of taxes, consisting of the FICA and federal income tax, while we also included the state income tax as well as the sales tax; lastly as already emphasized Feldstein employs a linearization procedure that ignores the issue of kink points while we take them into account. These three reasons are intertwined when explaining why the numerical results reported in Table 2 are not the same as those of Feldstein. The triangle approximation is known to lead to an underestimation of the deadweight loss. As this approximation is used by Feldstein, we have to slightly modify the set of taxes to replicate his numerical results. We do that by including a sales tax of 4.97%, corresponding to the population-weighted average across US states. Thanks to that, we get a value of the marginal deadweight loss per marginal tax dollar very close to that computed by Feldstein when we employ the linearization procedure as he does: we obtain a marginal deadweight loss per tax dollar of $2.16 when Feldstein finds a value of $2.06. In contrast, when calculating the marginal deadweight loss per tax dollar taking the curvature and kink points of the budget constraint into account, we find the marginal deadweight loss per tax dollar to be $1.35. This allows us to identify the effects of the linearization and to conclude that, in this situation, the linearization procedure overestimates the marginal deadweight loss per tax dollar by 61%. The different biases are shown in the next table.

Year 1994	DW DWL	TR TRL	DW/$	DWL/$
$\beta = 1.04$	$69 $82	$51 $38	1.35	2.16
Bias	+19%	−26%	+61%	

V.   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 first 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 next show how to calculate the marginal deadweight loss when the tax system generates a piecewise linear budget constraint. It is equally true in this case as for the case with a smooth budget constraint that the curvature of the budget constraint is of the same importance for the marginal deadweight loss as the curvature of the indifference curve. However, the impact of the curvature of the budget constraint to diminish the deadweight loss is now concentrated to the kink points. For individuals located at a kink point, there is no marginal deadweight loss, for them the increase in the marginal tax is just like a lump-sum tax.

We also perform numerical calculations where we calculate the true marginal deadweight loss and compare this with computations obtained by linearizing the budget constraint and performing the marginal deadweight calculations on the linearized budget constraint. The bias introduced by the linearization is often quite large, for reasonable parameter values.

It is very simple to use the correct procedure to compute the marginal deadweight loss. Therefore, there is no need to rely on a linearization procedure which leads to an incorrect measure.

References

Bastani, S. and H. Selin. 2011. "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule." UCFS Working Paper 2011:12.

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.

Blomquist, S. and H. Selin. 2010. "Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates." *Journal of Public Economics* 94:878–889.

Blundell, R. and H. Hoynes. 2004. *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000*, chap. Has In-Work Benefit Reform Helped the Labor Market? NBER, 411–460.

Blundell, Richard, Alan Duncan, and Costas Meghir. 1998. "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica* 66 (4):pp. 827–861.

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.

Burtless, G. and R. Moffitt. 1984. *Retirement and Economic Behavior*, chap. The Effect of Social Security Benefits on the Labor Supply of the Aged. Brookings Institution, 135–174.

Chapman, B. and A. Leigh. 2009. "Do Very High Tax Rates Induce Bunching? Implications for the Design of Income-Contingent Loan Schemes." *Economic Record* 85:276–289.

Chetty, R. 2009. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." NBER Working Paper 15616. Forthcoming in Econometrica.

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." *Annales des Ponts et Chausses* 2:332–375.

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.

Friedberg, L. 2000. "The Labor Supply Effects of the Social Security Earnings Test." *Review of Economics and Statistics* 82:48–63.

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.

Keane, Michael and Robert Moffitt. 1998. "A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review* 39 (3):pp. 553–589. URL http://www.jstor.org/stable/2527390.

Kimmel, Jean and Thomas J. Kniesner. 1998. "New evidence on labor supply:: Employment versus hours elasticities by sex and marital status." *Journal of Monetary Economics* 42 (2):289 – 301.

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

Kosters, M. 1967. *The Taxation of Income From Capital*, chap. Effects of an Income Tax on Labor Supply. Brookings Institution, 301–321.

MaCurdy, Thomas E. 1983. "A Simple Scheme for Estimating an Intertemporal Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty." *International Economic Review* 24 (2):pp. 265–289.

Prescott, Edward C. 2004. "Why do Americans Work so Much More than Europeans?" Working Paper 10316, National Bureau of Economic Research.

Rogerson, R. and J. Wallenius. 2009. "Micro and Macro Elasticities in a Life Cycle Model with Taxes." *Journal of Economic Theory* 144:2277–2292.

Saez, E. 2010. "Do Taxpayers Bunch at Kink Points." *American Economic Journal: Economic Policy* :forthcoming.

Saez, E., J. B. Slemrod, and S. H. Giertz. 2009. "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." NBER Working Paper 15012.