

Submission Number: PET11-11-00190

Marginal deadweight loss when the income tax is nonlinear

Laurent Simula
Uppsala University

Sören Blomquist
Uppsala University

Abstract

Almost all 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 the traditional linearization procedure is used. We perform calculations based on the 2006 US tax system and find that the relative deadweight loss caused by increasing existing tax rates is large but less than half of Feldstein's (1999) estimates for the 1994 tax system.

Marginal Deadweight Loss when the Income Tax is Nonlinear

Sören Blomquist and Laurent Simula

Uppsala University and Uppsala Center for Fiscal Studies*†

Jan. 25, 2010: Preliminary and Incomplete

Abstract

Almost all 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 the traditional linearization procedure is used. We perform calculations based on the 2006 US tax system and find that the relative deadweight loss caused by increasing existing tax rates is large but less than half of Feldstein's (1999) estimates for the 1994 tax system.

Keywords: Deadweight Loss, Taxable Income, Nonlinear Budget Constraint

JEL classification: H21, H24, H31, D61

*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 Thomas Gaube, Laurence Jacquet, Henrik Jacobsen Kleven, Olof Johansson Stenman, Etienne Lehmann, Luca Micheletto, Håkan Selin and Alain Trannoy for very helpful comments and suggestions.

1 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.

Almost all 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 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, one has linearized the nonlinear budget constraint and applied the procedure that is correct for variations in a linear income tax. 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 traditional linearization procedure is used. 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. We then describe how results are modified when the budget constraints are piecewise linear. 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.¹ To simplify the analysis, we consider tax systems that generate convex budget sets.² 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.

The layout of the rest of this article is as follows. Section 2 sets up the framework. Section 3 provides a simple example showing that linearizing a nonlinear budget constraint may lead to a significant mistake. Section 4 examines how serious it is to linearize in the general case. It shows that what matters is the relative curvature of the budget constraint in relation to the indifference curves and provides several ways of assessing the linearization bias. Section 5 uses empirical data and real tax systems, and quantifies the linearization bias in France, the US and Germany. Section 6 concludes.

2 The Framework

We start by considering the case with a smooth budget constraint for two reasons. One is simplicity. The basic idea comes through very clearly when the budget constraint is smooth and the analysis is simple. Another reason is that there are good arguments why we should analyze the budget constraint as if it were smooth, even if the statutory tax rules seem to imply a kinked budget constraint. As pointed out by Saez (2010), individuals cannot control their taxable income perfectly. Unforeseen bonus pay checks, better health than expected or assigned overtime would be examples of positive shocks. Unexpected sickness, a layoff, new extended vacation plans because of a new love would be examples of negative shocks. The individual does not know what kind of shock there will be, but he realizes that there is a random component in his taxable income. Hence, it is rational for the individual to take this random element into account when planning for his desired taxable income. Saez (2010) sets up a model for this. In this model individuals maximize their expected utility. Saez shows how this problem can be reformulated as a decision problem under certainty and a budget constraint that is smooth. According to this model, individuals behave as if they faced a smooth budget

¹This 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.

²This assumption is just for simplicity. The general insights of the article applies also to the case with a concave budget set. If the curvature of the indifference curves is larger than the curvature of the budget constraint so that an interior unique solution of the individual's utility maximization problem obtains, most of the formulas below apply. The analysis becomes more complex if the budget constraint is concave and more curved than the indifference curves. One then has to take the possibility of multiple solutions and/or corner solutions into account.

constraint although it in fact is kinked. What governs individuals' behavior is not the exact locations of the kink points, but the overall shape of the budget constraint. Of course, under this scenario the analysis with a smooth budget constraint is applicable.

2.1 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.

Therefore, we model the tax in the following way. Let A denote taxable income and the tax on A be given by $T(A)$. In the general case, 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 $g''(A)$. 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 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 chosen to parameterize the tax system in the way described above. When we vary the slope of a linear budget constraint, the intercept will not change. It is of value to have a parameterization of the nonlinear tax that has a similar property. When we in the next section study the marginal deadweight loss for a piecewise linear budget constraint, we will see that, for the parameterization used, a change in t will not change the virtual incomes but only the slope, thereby giving a clean 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 of a form as the one described by $g(A) + tA$.

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

2.2 Definition of the Marginal Deadweight Loss

Consider the utility maximization problem:

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

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 $A(t, B, v), C(t, B, v)$. The form of these two functions depends on the functional forms of U and g . Sticking $A(t, B, v), C(t, B, v)$ back into the utility function, we obtain the indirect utility $\bar{u}(v) := U(C(t, B, v), A(t, B, v), v)$. For each individual, the latter is the maximum utility level obtained under the given tax system. Because individuals have different v 's, they chose different taxable incomes and have different $\bar{u}(v)$. To simplify the notations, we henceforth suppress the v in $\bar{u}(v)$. However, it should be kept in mind that the $\bar{u}$ given in expressions below vary between individuals.

We now study the marginal deadweight loss of a small increase in t . We first derive the correct expression and then – in the next subsection – describe how it usually is calculated. For this purpose, we define the expenditure function as:

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

This problem also defines the compensated supply and demand functions, $A^h(t, v, \bar{u})$ and $C^h(t, v, \bar{u})$ 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.

Let us define the compensated revenue function as:

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

and the marginal deadweight loss as:

$$\begin{aligned} DW &:= \frac{dE(t, v, \bar{u})}{dt} - \frac{dR(A^h(t, v, \bar{u}))}{dt} \\ &= A^h - g'(A^h) \frac{dA^h}{dt} - A^h - t \frac{dA^h}{dt} = - \left(g'(A^h) + t \right) \frac{dA^h}{dt}, \end{aligned} \quad (2)$$

where we used the envelope theorem to obtain $dE(t, v, \bar{u})/dt = A^h$. Expression (2) is the correct expression for the marginal deadweight loss.

2.3 A Commonly Used Linearization Procedure

We next describe a commonly used procedure that, in general, overestimates the marginal deadweight loss. Let us consider particular values v^* , t^* and B^* and the solution to (P1), $A^* = A(t^*, v^*, B^*)$, $C^* = C(t^*, v^*, B^*)$. We can linearize the budget constraint around this point with local prices defined by $p_c = 1$ and $p_A = g'(A^*) + t^*$ 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:

$$\max_{A, C} U(C, A, v^*) \text{ s.t. } C \leq A - p_A A + M. \quad (\text{P3})$$

We call $A_L(p_A, v^*, M)$, $C_L(p_A, v^*, M)$ 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

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

and denote its solution by $A_L^h(t, v, \bar{u})$, $C_L^h(t, v, \bar{u})$, 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: $R(A_L^h(t, v, \bar{u})) = g(A_L^h(t, v, \bar{u})) + t A_L^h(t, v, \bar{u})$. We define the marginal deadweight loss as:

$$\begin{aligned} DW_L &:= \frac{dE_L(t, v, \bar{u})}{dt} - \frac{dR_L(A_L^h(t, v, \bar{u}))}{dt} \\ &= A_L^h - g'(A_L^h) \frac{dA_L^h}{dt} - A_L^h - t \frac{dA_L^h}{dt} = - \left(g'(A_L^h) + t \right) \frac{dA_L^h}{dt}. \end{aligned} \quad (3)$$

Figure 1 illustrates the links between the four problems that we have studied. The

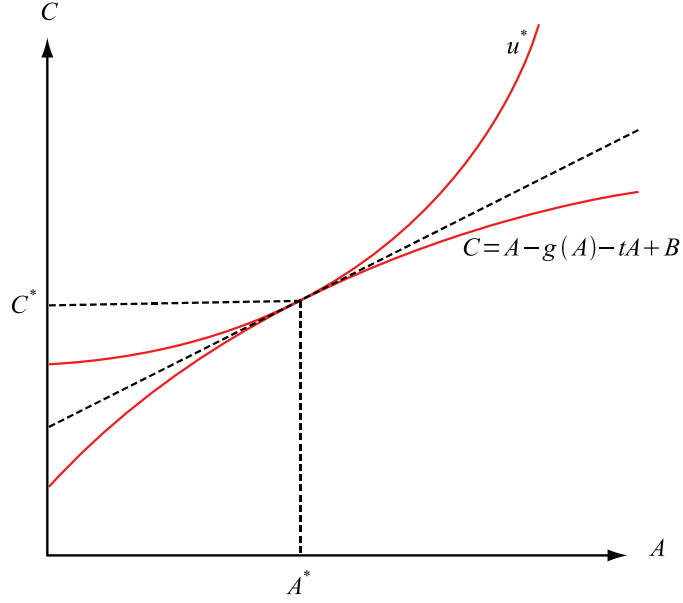


Figure 1: Nonlinear and Linearized Programmes

optimization problem (P1) maximizes utility given the curved budget constraint $C = A - g(A) - tA + B$ in the figure. 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 $A^* = A(t^*, B^*)$, $C^* = C(t^*, B^*)$. 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 A^*, C^* . 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

$$\begin{aligned} A(t, B) &\equiv A^h(U(C(t, B), A(t, B))) \\ &\equiv A_L(p_A(C(t, B), A(t, B)), M(C(t, B), A(t, B))) \equiv A_L^h(U(C(t, B), A(t, B))). \end{aligned} \quad (4)$$

Expressions (2) and (3) look quite similar. By construction, it is true that $A_L^h = A^h$, implying that $g'(A_L^h) + t = g'(A^h) + t$. However, dA^h/dt and dA_L^h/dt differ, implying

a bias when the linearization procedure is used. To show this, we start with a simple example, which we then generalize.

3 A Simple Example

To simplify notation, we in this example suppress the preference parameter v . Assume the utility function takes the quasilinear form $U = C - \alpha A - \beta A^2$. 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 a 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 $U = A - (p + t)A - \pi A^2 + B - \alpha A - \beta A^2$. Maximizing with respect to A , we get $dU/dA = 1 - (p + t) - 2\pi A - \alpha - 2\beta A$. We see that a necessary condition for a non-negative A is $1 - (p + t) - \alpha \geq 0$. We find that $d^2U/dA^2 = -2(\pi + \beta) < 0$ for $\pi + \beta > 0$. Setting $dU/dA = 0$ and solving for A , we obtain

$$A = \frac{1 - (p + t) - \alpha}{2(\pi + \beta)}. \quad (5)$$

Since we have the quasi-linear form, this is also the Hicksian supply. We immediately have

$$\frac{dA^h}{dt} = -\frac{1}{2(\pi + \beta)}. \quad (6)$$

From (6), 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 $2(\pi + \beta)$ in our example, 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 $M = C^* - [1 - (p + t) - 2\pi A^*]A^*$.

Consider the problem:

$$\max_{C, A} \{C - \alpha A - \beta A^2\} \text{ s.t. } C \leq A - [(p + t) + 2\pi A^*]A + M. \quad (7)$$

Substituting the binding budget constraint into the utility function, we want to maximise

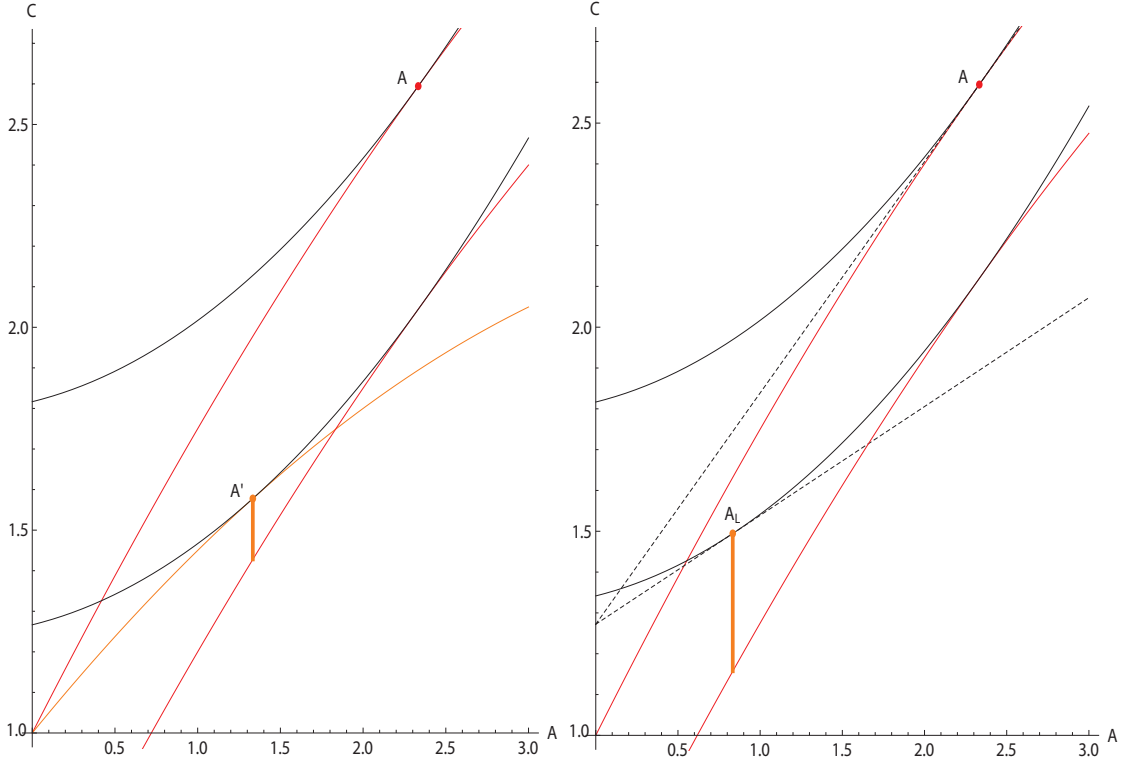


Figure 2: Deadweight loss when the budget constraint is nonlinear (left panel) and linearized (right panel)

$A - [(p + t) + 2\pi A^*] A + M - \alpha A - \beta A^2$. Denoting this expression by $\tilde{U}$, we obtain $d\tilde{U}/dA = 1 - (p + t) - 2\pi A^* - \alpha - 2\beta A$ and $d^2\tilde{U}/dA^2 = -2\beta$. The second-order condition is satisfied for $\beta > 0$. Setting $d\tilde{U}/dA = 0$ and solving for A , we get $A_L^h = (1 - (p + t) - 2\pi A^* - \alpha) / (2\beta)$ and

$$\frac{dA_L^h}{dt} = -\frac{1}{2\beta}. \quad (8)$$

Suppose $\pi = \beta = 0.1$. We then have that $dA^h/dt = -2.5$ 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.

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 $\alpha = \beta = 0.1$, $p = 0.2$, $\pi = 0.05$ and $B = 1$. In the left panel, we show the correct calculation of the deadweight loss 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. It is thus shown by the thick vertical line 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 line below A_L , is much larger than when the correct procedure is used.

4 How Serious is it to Linearize?

We want to know how serious it is to linearize to compute the marginal deadweight loss. To this aim, we can easily generalize the example above.

4.1 The Answer Depends on the Relative Curvature of the Indifference Curve and the Budget Constraint

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

$$\frac{dA^h}{dt} = -\frac{1}{g'' + f''}. \quad (9)$$

In the analysis above, $f'(A, v, \bar{u})$ is the slope of the indifference curve. Hence, $f''(A, v, \bar{u})$ 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, $\bar{u}$ would not be an argument in the $f()$ function. From (9), 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:

$$\frac{dA_L^h}{dt} = -\frac{1}{f''}, \quad (10)$$

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

$$\frac{dA^h}{dt} = \frac{dA_L^h/dt}{1 - g''(A)(dA_L^h/dt)}. \quad (11)$$

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. 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 .

4.2 Linearization Bias

The relative error in using the linearized budget constraint is given by the ratio of expressions (3) to (2), i.e. by:

$$\frac{dA_L^h/dt}{dA^h/dt} = \frac{g'' + f''}{f''} = 1 + \frac{g''}{f''} = 1 + a. \quad (12)$$

We see that the relative error in using the linearized budget constraint depends on the relative sizes of g'' and f'' . For simplicity, we call a the ratio g''/f'' . Then, a is a measure of the relative curvature of the budget constraint and the indifference curve. It is also a measure of the relative bias in the welfare measure if we incorrectly linearize. For example, if $a = 1$ and hence $g'' = f''$, the linearization procedure overstates the true effect by a factor 2. *This holds true irrespective of the absolute size of g'' and f'' .*

A high value of η_L , the elasticity with respect to a variation in the net tax rate, implying a low value of f'' and shallow indifference curves, gives a large bias even if

the curvature in the budget constraint is not very large. A low value of η_L (implying quite curved indifference curves) does not give a large bias unless the budget constraint also is heavily curved. An implication of the above is that for a given budget constraint the bias increases in η_L . So, the bias might be negligible for an elasticity around 0.2, which has been found in some studies (Blomquist and Selin, 2010) whereas it is large for an elasticity of 2, which is values that Prescott argues are correct. People looking at macro data find elasticities in the range of 2.25-3.0 (Rogerson and Wallenius, 2009). If we believe that the bias is getting serious if the bias is 20% or larger, this implies that if $a \geq 0.2$ the bias is serious.

How do we get from η_L to f'' ? Let $\theta = 1 - g' - t$, implying that $d\theta/dt = -1$. The taxable income elasticity is defined as

$$\eta_L := \frac{dA_L^h}{d\theta} \frac{\theta}{A} = -\frac{dA_L^h}{dt} \frac{\theta}{A} = \frac{1}{f''} \frac{\theta}{A}. \quad (13)$$

We can rewrite this to get $f'' = \frac{1}{\eta_L} \frac{\theta}{A}$.

We now examine how to get an idea about the size of g'' . We take the income interval $[D1, D9]$ ranging from the first to the ninth decile and consider the budget constraint for this range of income. We make the assumption that the second derivative of the tax function is constant over the interval. In other words, we approximate the nonlinear income tax by a quadratic function $g(A)$. We call t_{D1} and t_{D9} the marginal tax rates at the beginning and the end of the income interval, i.e., $t_{D1} \equiv g'(D1)$ and $t_{D9} \equiv g'(D9)$. Then, the function $g(A)$ can be written as

$$g(A) = \alpha + pA + g'' A^2/2 \text{ with } g'' = \frac{t_{D9} - t_{D1}}{D9 - D1}, \quad (14)$$

implying $\theta = 1 - (p + t) - Ag''$. This provides us with a very easy way to approximate the curvature of the tax function. Using this simple procedure, the bias at taxable income A amounts to

$$a := \frac{g''}{f''} = g'' \times \frac{\eta_L A}{1 - (p + t) - Ag''}. \quad (15)$$

4.3 How Large is the Linearization Bias? Two Illustrations

We can illustrate the above for the French tax system. We have chosen an income interval ranging from the first to the ninth decile (for men in the private and semi-public sectors) and computed the income tax using the Ministry of Finance's Website. In 2007, we have $D1 = 13.528$ K€, $D9 = 41.413$ K€, $t_{D1} = 19.14\%$ and $t_{D9} = 37.38\%$. We obtain

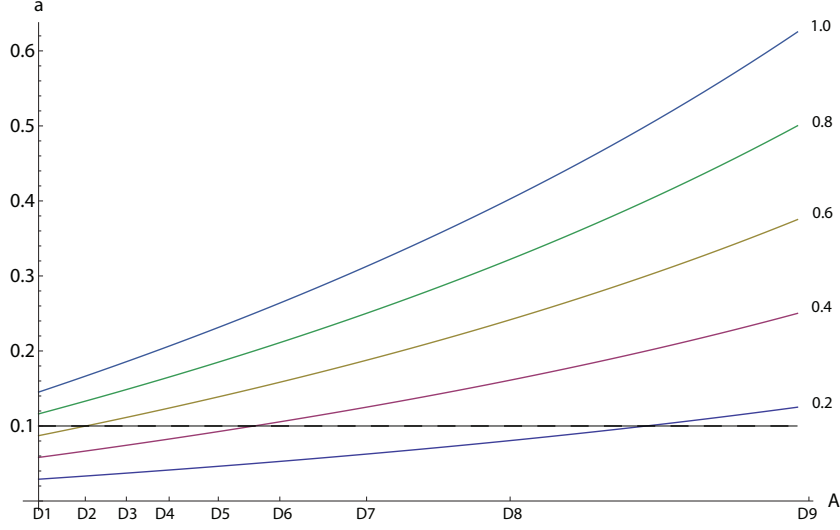


Figure 3: Bias a (single men, France, $t = 0.2$, 2007)

$p = 0.102911$ and $g'' = 0.0065412$. We assume that $t = 0.2$. Figure 3 shows the bias as a function of income for several elasticity values (0.2 to 1.0). D1, D2, ..., and D9 stand for the different deciles. We have checked that the bias is not very sensitive to the value of t . If we consider a 10%-mistake as significant, we see that the linearization procedure should be rejected for elasticities above 0.2–0.3.

We can also compute the average bias. If $\phi(A)$ is the proportion of individuals with taxable income A , the average mistake can be computed as:

$$\eta_L g'' \sum_{A=D1}^{D9} \frac{A}{1 - (p + t) - Ag''} \phi(A). \quad (16)$$

We provide computations for the US. We take into account the federal income tax, the state income tax, and the social security payroll tax. We use the Californian tax schedule to compute the state income tax. California is the state with the largest population and many other states have similar income tax schedules. The social security payroll tax is the linear component of the tax system we consider. We therefore investigate the deadweight loss induced by a marginal change in the social security tax. The distribution of taxable income is obtained from the CPS labor extracts, restricted to single men with no child, and the tax rates from TAXSIM. The different parameters are shown in Table 1. The marginal tax rates are shown as the sum of the federal (Fed.) and state (St.) income tax respectively.

	$D1$ (in \$)	$D9$ (in \$)	t_{D1} (Fed. + St.)	t_{D9} (Fed. + St.)	SS Tax	g''
1979	2,444	18,148	0%+0%	30%+11.1%	12.26%	0.0262
1993	4,940	33,020	0%+0%	28%+8%	15.3%	0.0128
2000	8,996	52,000	22.65%+0%	28%+9.3%	15.3%	0.0034
2007	10,400	62,400	17.65%+0%	25%+9.3%	15.3%	0.0032

Table 1: Descriptive Data for the US

We see that the overall curvature of the tax system – summarized by g'' – has dramatically changed over the last three decades. It was much larger in 1979 and 1993 than in the 2000s. Given g'' , we compute the bias a for each individual in our sample and obtain the average shown in Table 2. The reduction in the concavity of the federal and state income tax implies a reduction of the bias, from 14% in 1979 to 6% in 2007 for a low elasticity of 0.4 and from 35% to 16% for a larger elasticity of 1.0, that Feldstein argues as reasonable.

Year \ η_L	0.2	0.4	0.6	0.8	1.0	2.0
1979	0.07	0.14	0.21	0.28	0.35	0.71
1993	0.06	0.12	0.18	0.24	0.31	0.61
2000	0.03	0.06	0.09	0.12	0.16	0.31
2007	0.03	0.06	0.09	0.12	0.16	0.32

Table 2: Average Bias in the US

4.4 Marginal DWL per Marginal Tax Dollar

Sometimes one is interested in the marginal deadweight loss per marginal tax dollar. Because the compensated revenue function is $R(A^h) = g(A^h) + tA^h$, the marginal tax revenue – whilst keeping utility constant – can be written as:

$$\frac{dR(A^h)}{dt} = A^h + (g'(A^h) + t) \frac{dA^h}{dt} = A - \frac{g' + t}{g'' + f''}. \quad (17)$$

If a linearized budget constraint is used, the marginal compensated tax revenue will incorrectly be calculated as:

$$\frac{dR(A_L^h)}{dt} = A_L^h + (g'(A_L^h) + t) \frac{dA_L^h}{dt} = A - \frac{g' + t}{f''}. \quad (18)$$

That is, the marginal tax revenue will be *underestimated*. Hence, the linearization procedure does not only lead to an incorrect deadweight loss, it also computes a wrong change in tax revenue.

Let ρ denote the marginal deadweight loss per marginal tax dollar, i.e.,

$$\rho := \frac{\text{marginal deadweight loss}}{\text{marginal tax revenue}} = -\frac{(g'(A) + t) \, dA^h/dt}{dR(A^h)/dt} = \frac{\frac{g' + t}{g'' + f''}}{A - \frac{(g' + t)}{g'' + f''}}. \quad (19)$$

If we instead linearize, we would calculate:

$$\rho_L = \frac{\frac{g' + t}{f''}}{A - \frac{(g' + t)}{f''}}. \quad (20)$$

We find that $\rho_L/\rho = 1 + a_\rho$ with

$$a_\rho := \frac{Ag''}{Af'' - (g' + t)}, \quad (21)$$

which corresponds to the relative bias induced by the linearization procedure.

It is of interest to compare this bias with a . By definition, these biases are linked by the following relationship:

$$1 + a_\rho = \frac{DW_L}{DW} \times \frac{dR(A^h)/dt}{dR(A_L^h)/dt} = (1 + a) \times \frac{dR(A^h)/dt}{dR(A_L^h)/dt}, \quad (22)$$

that can also be rewritten as:

$$\frac{1 + a_\rho}{1 + a} = \frac{dR(A^h)/dt}{dR(A_L^h)/dt}. \quad (23)$$

We have already noted that $dR(A_L^h)/dt$ is larger than $dR(A^h)/dt$, i.e., $dR(A_L^h)/dt < dR(A^h)/dt$. When the marginal changes in tax revenue have the same sign (both positive or both negative), we get $a_\rho > a$. However, $dR(A_L^h)/dt$ and $dR(A^h)/dt$ can be of different signs. In that case, $a_\rho < a$.

Typically, $dR(A_L^h)/dt$ and $dR(A^h)/dt$ are both positive up to some income threshold. Then, for larger A , $dR(A_L^h)/dt$ becomes negative whilst $dR(A^h)/dt$ remains positive. The switching point corresponds to a vertical asymptote of a_ρ , as shown in Figure

4: the bias explodes for incomes close to it, in the range of positive numbers to the left and of negative numbers to the right. This is due to the fact that the linearization procedure can suggest that tax revenue is decreasing (with respect to t) even though it is not yet. In other words, *computations ignoring the curvature of the budget constraint can show that we are on the left side of the Laffer curve even though we are not.*

Figure 4 is obtained for $\eta_L = 1$. For smaller elasticity values, the asymptot is more to the right, but the overall pattern is the same. Showing the graph of the bias for several elasticity values on the same figure would be quite messy because of the different asymptots. This is why, in Figure 5, we have chosen to focus on the positive part of the plane.

Year \ η_L	0.2	0.4	0.6	0.8	1.0	2.0
1979	0.08	0.18	0.34	0.66	0.43	1.71
1993	0.07	0.16	0.18	0.52	2.42	1.47
2000	0.04	0.10	0.20	0.44	1.87	-0.55
2007	0.04	0.09	0.17	0.33	1.01	-1.89

Table 3: Average Bias a_ρ in the US

5 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

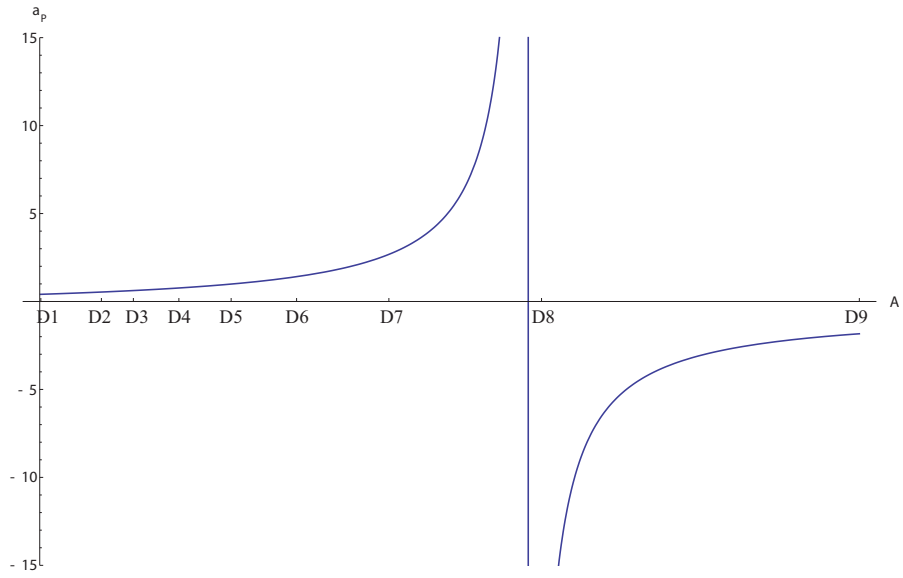


Figure 4: Relative Bias a_ρ for $\eta_L = 1$ (single men, France, $t = 0.2$, 2007)

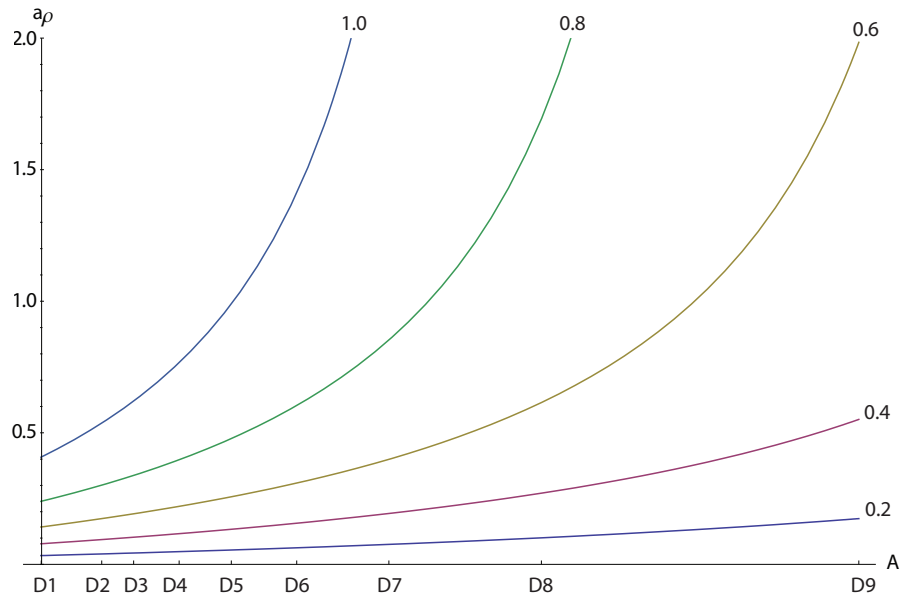


Figure 5: Relative Bias a_ρ in the Positive Plane (France, 2007, with $t = 0.2$)

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.

6 Appendix

Desired, or planned, taxable income is determined as outlined above. However, for various reasons desired taxable income cannot be realized. To take this fact into account, δ introduces a random shock δ , corresponding to an increase or decrease in labour earnings. Wage bonus is an example of positive shock. We can think of other reasons. The individual might plan for a given taxable income and choose his effort/labour supply accordingly. 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. Taxable income might be higher than the income planned for because of vacation plans that are changed, better health than expected, assigned overtime, etc. We call ε this kind of shocks. We assume that ε and δ are independent, with supports $[\underline{\varepsilon}, \bar{\varepsilon}]$ and $[\underline{\delta}, \bar{\delta}]$ and pdf

E and D respectively. They both have means equal to zero.

Assuming that there is no income effect on taxable income, the utility is given by:

$$U = A + \varepsilon + \delta - T(A + \varepsilon + \delta) - v \left(\frac{A + \varepsilon}{w} \right) \quad (24)$$

where v is disutility of effort/labour. The expected utility is:

$$\begin{aligned} EU &= \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \int_{\underline{\delta}}^{\bar{\delta}} \left[A + \varepsilon + \delta - T(A + \varepsilon + \delta) - v \left(\frac{A + \varepsilon}{w} \right) \right] E(\varepsilon) D(\delta) d\varepsilon d\delta \\ &= A - \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \int_{\underline{\delta}}^{\bar{\delta}} T(A + \varepsilon + \delta) E(\varepsilon) D(\delta) d\varepsilon d\delta - \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} v \left(\frac{A + \varepsilon}{w} \right) E(\varepsilon) d\varepsilon. \end{aligned} \quad (25)$$

We call $\hat{T}(A) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \int_{\underline{\delta}}^{\bar{\delta}} T(A + \varepsilon + \delta) \times E(\varepsilon) D(\delta) d\varepsilon d\delta$ the expected tax and $\hat{T}'(A) = d\hat{T}(A)/dA$ the expected marginal tax rate. The first-order condition of the utility maximisation programme yields:

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} v' \left(\frac{A + \varepsilon}{w} \right) \times E(\varepsilon) d\varepsilon = w \left[1 - \hat{T}'(A) \right]. \quad (26)$$

We see that the optimum A depends on the distributions and supports of ε and δ as well as on the expected net-of-tax wage rate. In order to make their choice, individuals do not consider the actual tax schedule, but the expected smooth one.

References

- 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.
- 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.

- 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.
- KOPCZUK, W. (2005): “Tax bases, tax rates and the elasticity of reported income,” *Journal of Public Economics*, 89, 2093–2119.
- ROGERSON, R., AND J. WALLENIS (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*, p. 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.