



Marginal deadweight loss when the income tax is nonlinear

Sören Blomquist^{a,*}, Laurent Simula^b

^a Uppsala Center for Fiscal Studies, Uppsala University, Department of Economics, P.O. Box 513, SE-75120 Uppsala, Sweden

^b Ecole Normale Supérieure de Lyon and UMR GATE-LSE, BP 7000, 69342 Lyon cedex 07, France

ARTICLE INFO

Article history:

Available online 18 December 2018

JEL classification:

H21
H24
H31
D61

Keywords:

Marginal deadweight loss
Taxable income
Nonlinear budget constraint

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 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, 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. Our main theoretical result is that the overall curvature of the tax system plays the same role as the curvature of indifference curves for the size of the marginal deadweight loss. Using numerical simulations calibrated on US data, we show that common linearization procedures may lead to substantial overestimation of the marginal deadweight loss.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The study of the deadweight loss (or excess burden) of taxation has a long tradition in economics going back as far as Dupuit (1844). Modern type of empirical work on the deadweight loss of taxation is heavily influenced by the important work of Harberger in the 1950s and 1960s (see for example Harberger, 1962, 1964). A second generation of empirical work was inspired by Feldstein (1995, 1999). The latter 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 et al., 2012). These taxable income functions show how taxable income varies as the slope of a linear budget constraint 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 (see Auerbach and Hines, 2002, for an excellent survey). This is quite surprising because most income tax systems are nonlinear, generating nonlinear budget constraints. Instead of developing a procedure to calculate the marginal deadweight loss for variations in nonlinear income taxes, one usually linearizes the nonlinear

* Corresponding author.

E-mail addresses: Soren.Blomquist@nek.uu.se (S. Blomquist), Laurent.Simula@ens-lyon.fr (L. Simula).

budget constraint and applies a procedure that is correct for variations in a linear income tax. In empirical work estimating behavioral parameters, it is also common to linearize budget constraints and pursue the analysis with the linearized budget constraints. However, when doing this, one is well aware that this linearization creates econometric problems and there is a large literature on methods to handle them (see e.g., [Burtless and Hausman, 1978](#); [Hausman, 1985](#); [Blomquist and Newey, 2002](#)). There is no similar literature on the problems that arise when one linearizes budget constraints and uses the linearized budget constraints to do welfare analysis. The main purpose of our article is therefore 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.

Although actual tax systems are usually piecewise linear, we begin our theoretical analysis by focusing on a smooth budget constraint. A first reason is that we then get very simple and clean results. A second reason is that agents may behave as if they are facing a smooth budget constraint, even though the statutory tax schedule is piecewise linear. As emphasized in early studies in the nonlinear budget set literature (e.g., [Burtless and Hausman, 1978](#); [Hausman, 1979, 1985](#); [Blomquist, 1983](#)) and more recently by [Chetty \(2012\)](#), there usually is a difference between desired taxable income (hours of work) and realized taxable income because of optimization errors. Under certain assumptions, [Saez \(1999\)](#) shows that if individuals are aware of the fact that there are optimization errors and take this into account when maximizing expected utility, this amounts to utility maximization subject to a smooth budget constraint. We then show how the analysis is modified when the budget constraint is piecewise linear. It however should be noted that the average, or aggregate, behavior for a population should not significantly depend on whether the tax system is kinked or smooth, provided the distribution of the kinks is regular enough. It is indeed the general shape of the budget constraints that determine the average behavior.

We use the definition of the marginal deadweight loss building on the equivalent variation; however our contribution remains valid for alternative definitions. To simplify the theoretical analysis, we consider tax systems that generate convex budget sets. The analysis can easily be modified to account for non-convexities, and we fully account for the latter in our numerical simulations. In addition, we state our results in terms of taxable income rather than labor supply, because the most recent literature focuses on this concept. It is easy to modify our results to some other application. To be as general as possible, we investigate the effects of a change in an arbitrary parameter of the tax function. To make our analysis more transparent, we then pay special attention to a change in the tax schedule such that the individual budget constraints rotate downwards and the marginal tax rate increases with the same number of percentage points at all income levels. This change is easily comparable with a tilt in a linear budget constraint and can, for example, be interpreted as an increase in the payroll tax. We allow for individual heterogeneity, which plays an important part in quantitative exercises.¹ There exists an important literature studying different ways of aggregating deadweight losses. We however make no contribution to this literature, the research question of which is orthogonal to the main point we make. We chose to report the population-weighted average deadweight loss in our simulations, as this measure is often used, but our main point remains valid for alternative weights as will become clear later on.

Our main result can be stated as follows. It is fairly well known that the marginal deadweight loss is proportional to the compensated change in taxable income. It is also well understood that the compensated change is smaller the more curved the indifference curves are. Our principal contribution is that, when the budget constraint is nonlinear, due to a nonlinear income tax, the curvature of the budget constraint plays the same role for the size of the compensated change as the curvature of the indifference curve: the more curved the budget constraint, the smaller the compensated change. The problem with linearizing is that one sets the curvature of the budget constraint to zero and thereby overestimates the size of the compensated change. For tax systems where the marginal income tax increases with the taxable income, the marginal deadweight loss obtained when linearizing is larger than the actual one.

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 the marginal deadweight loss and underestimates the change in the 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, and the bias in this measure may become very large.

As will be shown below, the expression for the marginal deadweight loss when a linearization procedure is used looks, at least for some nonlinear tax functions, formally the same as the correct expression. The crucial difference is that when the linearization procedure is used, one misses the fact that the change in taxable income depends on the curvature of the budget constraint. By linearizing, one sets the curvature of the budget constraint to zero and overestimates the change in taxable income.

Our article is organized as follows. It is simplest to illustrate the basic ideas under the assumption that the budget constraints are smooth. In Section 2, 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 3, we

¹ [Burtless and Hausman \(1978\)](#) showed the importance of allowing for individual heterogeneity when studying labor supply. [Hausman and Newey \(1995, 2016\)](#) show the importance of allowing for individual heterogeneity for doing welfare analysis. [Dahlby \(1998\)](#), who studies the marginal cost of public funds, also emphasizes the importance of allowing for heterogeneity.

show how the calculations are modified if budget constraints are piecewise linear. To illustrate the empirical importance of the curvature of the indifference curves and the budget constraint we in Section 4 present calculations of the marginal deadweight loss using the US tax systems in 1979, 1993 and 2006; which vary with respect to the curvature of the implied budget constraints. To vary the curvature of the indifference curves we use a convenient parametric form for the utility function where the curvature and taxable income elasticity depends on a single parameter in the utility function. To perform the calculations we also have to specify a distribution of preferences. Section 5 concludes.

2. 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. 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 many different ways. We perform our analysis for an arbitrary tax parameter. The main insight of our analysis will be that the comparative statics for the compensated taxable income depends on the curvature of the budget constraint in the same way as it depends on the curvature of the indifference curve.

Consider a tax function $T(A, \theta)$ where A is taxable income and θ is a vector of tax parameters. Without loss of generality, to simplify notation, we consider an arbitrary parameter in this vector and denote it by γ . In the rest of this section, we suppress the other parameters in the tax function. We assume $\partial T(A, \gamma) / \partial A > 0$, $\partial^2 T(A, \gamma) / \partial A^2 > 0$, $\partial T(A, \gamma) / \partial \gamma > 0$ and $\partial^2 T(A, \gamma) / \partial A \partial \gamma > 0$.

2.2. The marginal deadweight loss

Consider the utility maximization problem:

$$\text{Max}_{C,A} U(C, A, v) \quad \text{s.t.} \quad C = A - T(A, \gamma) + B \tag{P1}$$

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 $C(\gamma, B, v)$, $A(\gamma, B, v)$. The form of these two functions depends on the functional forms of U and T . Sticking $C(\gamma, B, v)$, $A(\gamma, B, v)$ back into the utility function, we obtain the indirect utility $\bar{u}(v) := U(C(\gamma, B, v), A(\gamma, B, v), v)$. For each individual, the latter is the maximum utility level obtained under the given tax system. Because individuals have different v , they choose different taxable incomes and have different $\bar{u}(v)$. To save on notations, we below often drop the v and simply write \bar{u} .

We now study the marginal deadweight loss of a small increase in the tax parameter γ . We also examine the marginal deadweight loss per marginal tax dollar, as this measure is often used in the literature. For this purpose, we define the expenditure function as:

$$E(\gamma, v, \bar{u}) = \text{Min}_{C,A} \{C - A + T(A, \gamma) - B\} \quad \text{s.t.} \quad U(C, A, v) \geq \bar{u}. \tag{P2}$$

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

$$T(A^h(\gamma, v, \bar{u}), \gamma) \tag{1}$$

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

$$MTR = \frac{\partial T(A^h, \gamma)}{\partial A} \frac{\partial A^h(\gamma, v, \bar{u})}{\partial \gamma} + \frac{\partial T(A^h, \gamma)}{\partial \gamma}. \tag{2}$$

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

$$MDW = \frac{dE(\gamma, v, \bar{u})}{d\gamma} - MTR = \frac{\partial T(A^h, \gamma)}{\partial \gamma} - \frac{\partial T(A^h, \gamma)}{\partial A} \frac{\partial A^h(\gamma, v, \bar{u})}{\partial \gamma} - \frac{\partial T(A^h, \gamma)}{\partial \gamma} = \frac{\partial T(A^h, \gamma)}{\partial A} \frac{\partial A^h(\gamma, v, \bar{u})}{\partial \gamma} \tag{3}$$

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

2.3. Comparative statics for taxable income

From Eq. (3), we see that the compensated change $\partial A^h/\partial \gamma$ is key for the size of the marginal deadweight loss. In this subsection, we therefore study the determinants of this compensated change.

Assume the constraint $U(C, A, v) \geq \bar{u}$ is binding. We can then define the function $C = f(A, v, \bar{u})$. Let $f'(A, v, \bar{u}) = \partial f/\partial A$ and $f''(A, v, \bar{u}) = \partial^2 f/\partial A^2$. We note that f' is the slope of the indifference curve with utility level \bar{u} and f'' the curvature of the indifference curve. Substitute $C = f(A, v, \bar{u})$ into the objective function in problem (P2) to obtain the minimization problem: $\text{Min}_A f(A, v, \bar{u}) - A + T(A, \gamma) - B$. Using the notations $T'(A, \gamma) = \partial T/\partial A$ and $T''(A, \gamma) = \partial^2 T/\partial A^2$, we obtain the first order condition $f'(A) - 1 + T'(A, \gamma) = 0$, which defines the Hicksian taxable income supply function. Differentiating implicitly with respect to γ , we obtain:

$$f''(A) \frac{\partial A^h}{\partial \gamma} + T''(A) \frac{\partial A^h}{\partial \gamma} + \frac{\partial^2 T}{\partial A \partial \gamma} = 0,$$

which can be rewritten as:

$$\frac{\partial A^h}{\partial \gamma} = - \frac{\partial^2 T/\partial A \partial \gamma}{T'' + f''} \quad (4)$$

The expression for the marginal deadweight loss (MDW) is therefore:

$$MDW = - \frac{\partial T}{\partial A} \frac{\partial A^h}{\partial \gamma} = \left(\frac{\partial T}{\partial A} \right) \left(\frac{\partial^2 T/\partial A \partial \gamma}{T'' + f''} \right) \quad (5)$$

The budget constraint can be written as $C = A - T(A, \gamma)$. Its slope is therefore given by $dC/dA = 1 - T'(A)$ and its curvature by $d^2C/dA^2 = -T''(A)$. From Eq. (5), we see that the curvatures of the indifference curve and the absolute value of the curvature of the budget constraint (equal to the curvature of the tax function) are of equal importance for the size of the deadweight loss. This result holds for a change in *any* tax parameter.

2.4. Linearization

In the past, when researchers have calculated the marginal deadweight loss of an increase in a nonlinear tax, they have not used the formulas derived above. Instead they have linearized the budget constraint and calculated the marginal deadweight loss based on formulas valid for linear budget constraints; as we will show this can lead to a serious bias in the marginal deadweight loss calculation. Feldstein (1999) is an early example of such a linearization. Kleven and Kreiner (2006), Eissa et al. (2008) and Gelber and Mitchell (2012) are examples of recent prominent articles that employ a linearized budget constraint to calculate deadweight losses.²

The linearization is usually done in the following way. Suppose an individual has chosen a point on his budget constraint such that taxable income and consumption are A^* , C^* . One then represents the tax system for this individual with a proportional tax $\tau = T'(A^*, \gamma)$. For the linear tax system to generate the actual values A^* , C^* , the lump-sum tax component must be defined as $T_0 = A^* - C^* - \tau A^* + B$, generating the budget constraint $C = A - T_0 - \tau A + B$. We now consider the problem:

$$\text{Max}_{C,A} U(C, A, v) \quad \text{s.t.} \quad C = A - T_0 - \tau A + B \quad (P3)$$

We call the Marshallian functions $C_L(\tau, v, B - T_0)$ and $A_L(\tau, v, B - T_0)$ the solution to this problem. Here, we use the subscript L to show that these are functions generated by a linear budget constraint. For τ and T_0 as defined above, the solution will be equal to the actual values A^* , C^* .

Applying the results obtained in Section 2.2 for a general non-linear tax function to the case with a linearized tax function we derive the Hicksian taxable income function $A_L^h(\tau, v, \bar{u})$ and write the marginal tax revenue, keeping utility constant, as:

$$MTR_L := A_L^h(\tau, v, \bar{u}) + \tau \frac{\partial A_L^h(\tau, v, \bar{u})}{\partial \tau} \quad (6)$$

and the marginal deadweight loss as:

$$MDW_L := -\tau \frac{\partial A_L^h(\tau, v, \bar{u})}{\partial \tau} \quad (7)$$

² Kleven and Kreiner (2006) extend the theory and measurement of the marginal cost of public funds to account for labor force participation responses. Eissa et al. (2008) embed the participation margin in an explicit welfare theoretical framework and show that not modeling the participation decision induces large errors. Gelber and Mitchell (2012) examine how income taxes affect time allocation during the entire day, and how these time allocation decisions interact with expenditure patterns.

The simplest way to obtain $\partial A_L^h / \partial \tau$ is to use Eq. (4) but realizing that $A_L^h(\tau, v, \bar{u})$ is generated from a linear budget constraint with zero curvature. This gives:

$$\frac{\partial A_L^h(\tau, v, \bar{u})}{\partial \tau} = -\frac{\partial^2 T / \partial A \partial \tau}{f''} = -\frac{1}{f''}$$

Eqs. (6) and (7) respectively give the marginal tax revenue and the marginal deadweight loss of a variation in τ , but what we are interested in is the marginal tax revenue and marginal deadweight loss of a variation in γ . It is not totally clear how this should be done. Here we describe the procedure used in Eissa et al. (2008).³ Translating their notation into the notation used in this article, they define the compensated change as:

$$\frac{\partial A_L^h}{\partial \gamma} = \frac{\partial A_L^h(\tau, v, \bar{u})}{\partial \tau} \frac{\partial^2 T}{\partial A \partial \gamma}$$

and the marginal deadweight loss as

$$MDW_L = -\frac{\partial T(A, \gamma)}{\partial A} \frac{\partial A_L^h(\tau, v, \bar{u})}{\partial \tau} \frac{\partial^2 T}{\partial A \partial \gamma} = \frac{\partial T(A, \gamma)}{\partial A} \frac{1}{f''} \frac{\partial^2 T}{\partial A \partial \gamma} \quad (8)$$

2.5. Bias when linearizing

We have chosen to study the bias that results when a linearization procedure is used for a simple but still important tax change. In a stylized way, we can write the US tax system as

$$T(A, t) = g(A) + tA \quad (9)$$

where the function $g(A)$ can be thought of as a general nonlinear tax function and tA as a proportional 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. We will study the marginal deadweight loss of a marginal increase in t .

For the tax function (9), the marginal tax revenue function (2) reduces to:

$$MTR = A^h(t, v, \bar{u}) - (g'(A^h(t, v, \bar{u})) + t) \left(\frac{1}{g'' + f''} \right) \quad (10)$$

and the marginal deadweight loss function (5) to:

$$MDW = (g'(A^h(t, v, \bar{u})) + t) \left(\frac{1}{g'' + f''} \right) \quad (11)$$

The expressions which build on the linearization procedure used by Eissa et al. (2008) reduce to

$$MTR_L = A_L^h(t, v, \bar{u}) - (g'(A_L^h(t, v, \bar{u})) + t) \frac{1}{f''} \quad (12)$$

and

$$MDW_L = (g'(A_L^h(t, v, \bar{u})) + t) \frac{1}{f''} \quad (13)$$

Eqs. (12) and (13) look quite similar to (10) and (11). By construction of the linearization, $A_L^h = A^h$. However, $\partial A_L^h / \partial t = -1/f''$ while $\partial A^h / \partial t = -1/(g'' + f'')$. Consequently, linearizing leads to biased measures of the marginal tax revenue and of the marginal deadweight loss.

We will measure the bias in three ways: overestimation of the marginal deadweight loss, underestimation of the marginal tax revenue and overestimation of the marginal deadweight loss per marginal tax dollar. 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:

$$\frac{MDW_L - MDW}{MDW} = \frac{g''}{f''} = a \quad (14)$$

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

³ Eissa et al. (2008) consider changes both along the extensive and intensive margins. The equation we give here is their expression for the compensated change along the intensive margin.

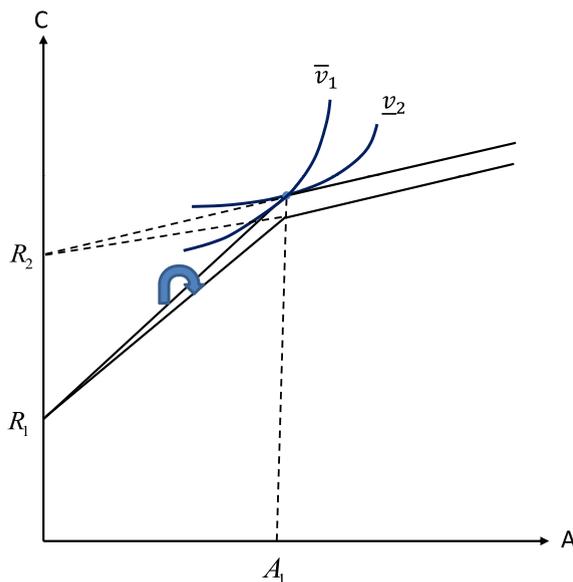


Fig. 1. Piecewise Linear Budget Constraint.

The relative curvature of the budget constraint and the indifference curve as measured by a also plays an important role in the expression for the relative error in marginal tax revenues. However, f'' and g'' enter this expression in other ways. The relative error in the marginal tax revenue is obtained as:

$$\frac{MTR_L - MTR}{MTR} = a \frac{(g'(A^h) + t) \partial A_L^h / \partial t}{A^h + (g'(A^h) + t) \partial A^h / \partial t} \tag{15}$$

The bias in marginal deadweight loss per marginal tax dollar is given by:

$$\frac{MDW_L / MTR_L - MDW / MTR}{MDW / MTR} = \frac{g'' A^h}{f'' A^h - (g'(A^h) + t)} \tag{16}$$

Once again, we see the importance of the curvature of the budget constraint.

3. 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. To keep the analysis straightforward, we consider a tax system of the same form as in (9), with $T(A) = g(A) + tA$ but $g(A)$ is now piecewise linear. In this setting, we contemplate the effects of a change in t on deadweight loss measures.⁴

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. More specifically, we consider that $g(A)$ 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. The corresponding budget constraint is shown in Fig. 1. The slope is given by $\theta_1 = 1 - \tau_1 - t$ on the first segment and by $\theta_2 = 1 - \tau_2 - t$ on the second segment. 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.

3.1. 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 explicitly re-introduce the heterogeneity preference parameter v and write the utility function $U(C, A, v)$, where v is a preference parameter with pdf $\varphi(v)$ over (\underline{v}, \bar{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, because in the taxable income

⁴ A referee suggested that the piece-wise linear budget constraint maybe could be regarded as a special case of the continuous one, where curvature is infinite at some points. We do not pursue that idea in the present article but leave it for future research.

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.

Agents with $A_L(\tau_1 + t, v, R_1) > A_1 > A_L(\tau_2 + t, v, R_2)$ are infra-marginal with respect to the kink. Given that the supply function $A_L(t, v, R)$ is continuous, all of them still locate at the kink following the infinitesimal tax change. They do not have tangency solutions and we can therefore not discuss the marginal deadweight loss for them in terms of income and substitution effects, but must fall back on the basic definition of deadweight loss. Let us denote the consumption before the tax change by \bar{C} . The utility before the tax change is therefore given by $U(\bar{C}, A_1, v)$. Utility after the tax change is $U(\bar{C} - dt \times A_1, A_1, v)$. Going back to the basic definition of deadweight loss, we ask how much money has to be transferred in a lump-sum way to restore one of these agents' utility to its original level. The amount needed is $dt \times A_1$, which exactly corresponds to the amount of tax collected via the tax change. Hence there is no deadweight loss of the tax change for an individual located at the kink both before and after the tax change.

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. The marginal deadweight loss for an individual with parameter v would be given by

$$-(\tau_i + t) \frac{\partial A_L^h(\tau_i + t, v, \bar{u})}{\partial t},$$

with $i = 1, 2$, where \bar{u} is indirect utility of a v -individual under the initial budget constraint.

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 $MTR = A + (\tau_i + t) \partial A / \partial t$ whereas we obtain $MTR = A$ for a person with desired taxable income at the kink point.

3.2. Marginal deadweight loss for the population

We consider that the preference parameter is continuously distributed over (v, \bar{v}) , with pdf $\varphi(v)$. If we want to compute the aggregate marginal deadweight loss, we therefore can integrate over v . For simplicity, we assume that v enters the utility function in such a way that A_L and A_L^h are strictly increasing in v . We also assume that $0 < A_L(\tau_1 + t, v, R_1) < A_1 < A_L(\tau_2 + t, v, R_2)$.⁵ 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 \bar{v}_1 be defined by $A_L(\tau_1 + t, \bar{v}_1, R_1) = A_1$ and v_2 by $A_L(\tau_2 + t, \bar{v}_1, R_2) = A_1$ as shown in Fig. 1. Define the subsets $S_1 = (v, \bar{v}_1)$ and $S_2 = (v_2, \bar{v})$. Likewise define the set $K_1 = (\bar{v}_1, v_2)$. Then individuals with v in S_1 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.

Consider a small finite increase dt in the tax t . Because $A_L(\tau_i + t, \bar{v}_i, R_i) = A_1$, implicit differentiation yields $\frac{\partial \bar{v}_i}{\partial t} = -\frac{\partial A_L / \partial (\tau_i + t)}{\partial A_L / \partial \bar{v}_i} > 0$ for $i = \{1, 2\}$. The interval of v for which there is an optimum at the kink moves to the right. Denoting the new interval by (v_1, v_2) , we have: $\bar{v}_1 < v_1 < v_2 < v_2$. The population can therefore be divided into five groups:

- For $v < \bar{v}_1$, there is a tangency solution both before and after the tax change and marginal deadweight loss is given by formula (7). We denote it by $MDW_L(v)$.
- For $v \in (\bar{v}_1, v_1)$, agents had a solution at the kink from which they move out to have a tangency solution along the first segment. The marginal deadweight loss is now greater than zero but below that given by formula (7). We denote it by $MDW(v)$.
- For $v \in (v_1, v_2)$, the optimum is at the kink both before and after tax change. Marginal deadweight loss for these persons is zero.
- For $v \in (v_2, v_2)$, agents had a solution along the second linear segment and now locate at the kink. The marginal deadweight loss for them, denoted by $MDW(v)$, is greater than zero but less than that given by formula (7).
- For $v > v_2$, here is a tangency solution both before and after tax change, so that the deadweight loss is given by $MDW_L(v)$.

Consequently, the aggregate marginal deadweight loss of an increase dt is the sum of five integrals:

$$\int_{v < \bar{v}_1} MDW_L(v) \varphi(v) dv + \int_{\bar{v}_1}^{v_1} MDW(v) \varphi(v) dv + \int_{v_1}^{v_2} 0 \times \varphi(v) dv + \int_{v_2}^{v_2} MDW(v) \varphi(v) dv + \int_{v > v_2} MDW_L(v) \varphi(v) dv \tag{17}$$

⁵ The assumption of monotonicity is not needed for our results. We only make the assumption that A_L and A_L^h are strictly increasing in v to simplify the exposition.

The tricky parts are the second and fourth terms. By definition of \bar{v}_1 and \bar{v}_2 , we can write:

$$v_i = \bar{v}_i + d\bar{v}_i = \bar{v}_i - \frac{\frac{\partial A_L}{\partial(\tau_i+t)}}{\frac{\partial A_L}{\partial \bar{v}_i}} dt \quad \text{for } i = \{1, 2\}$$

where the partial derivatives are obtained by the implicit differentiation. It thus follows that v_i tends to \bar{v}_i when dt tends to zero. This implies that the mass in the second and fourth integrals of (17) go to zero. The aggregate marginal deadweight loss of a marginal increase dt in t is thus given by:

$$\begin{aligned} ADWL &= \int_{v < \bar{v}_1} MDWL_L(v) \varphi(v) dv + \int_{v > \bar{v}_2} MDWL_L(v) \varphi(v) dv \\ &= - \int_{v < \bar{v}_1} (\tau_i + t) \frac{\partial A_L^h}{\partial t} \varphi(v) dv - \int_{v > \bar{v}_2} (\tau_i + t) \frac{\partial A_L^h}{\partial t} \varphi(v) dv \end{aligned} \quad (18)$$

Applying the same reasoning to the aggregate marginal tax revenue, we obtain:

$$\begin{aligned} AMTR &= \int_{v < \bar{v}_1} \left[A + (\tau_i + t) \frac{\partial A_L^h}{\partial t} \right] \varphi(v) dv + \int_{v_1}^{v_2} A \varphi(v) dv \\ &+ \int_{v > \bar{v}_2} \left[A + (\tau_i + t) \frac{\partial A_L^h}{\partial t} \right] \varphi(v) dv \end{aligned}$$

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 for the case with a smooth budget constraint or, when the budget constraint is piece-wise linear, it is neglected that desired taxable income will be at a kink point for part of the population.⁶ 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.

4. Numerical Computations

In this section, we investigate the empirical importance of using the correct procedure for calculating the marginal deadweight loss. We do not look at the impact of a general tax parameter, but consider the tax function (9), with $g(A)$ being piece-wise linear, and the impact of a marginal change in t . The analysis in Sections 3 and 4 has been for general utility functions and distributions of preferences. However, when we in this section compute the deadweight loss using actual data we have to make parametric assumptions about the utility function and the distribution of preferences. As will be shown we also must make assumptions about the distribution of optimization errors. Our objective is to illustrate the importance of accounting for the curvature of the budget constraint. The results obtained will, of course, depend on the parametric assumptions made. Hausman and Newey (1995, 2016) show how to do nonparametric welfare analysis when budget constraints are linear. It is an interesting future research task to see if the results of the present paper can be extended to a nonparametric welfare analysis.

We calculate the marginal deadweight loss for the US tax systems in 1979, 1993 and 2006. We choose three different years as the shape of the tax schedule, and, thus, the curvature of the implied budget constraints, have changed over time. Note that 1993 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 seems to be very little bunching at kinks of the US tax schedule.

4.1. 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 1960–1997, Saez (2010) finds clear evidence of

⁶ In an important article, Dahlby (1998) studies how to measure the marginal cost of public funds for variations in a piecewise linear tax system when the population is heterogeneous. He provides several alternative measures and compares them. However, he implicitly assumes that all desired taxable incomes are in the interior of linear segments. Hence, his formulas do not take account of kink points. Using the procedure presented above, it is easy to incorporate the part played by kinks into Dahlby's formulas (see also Dahlby, 2008).

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 (2014) study the Swedish tax system and 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) and Hausman (1985) have emphasized that there is usually a difference between desired taxable income (hours of work) and realized taxable income (hours of work) 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. More recently, Chetty (2012) argues that because of optimization frictions, we would not observe much bunching at kink points even if desired hours of work (or taxable income) are at a kink point. He computes the gains from bunching at kinks in the 2006 US tax schedule, instead of optimizing under the incorrect assumption that the tax rate in the previous bracket continues into the next bracket. He finds that the utility losses are less than 1% of net earnings at most kinks. It is optimal for agents to deviate from their frictionless optima provided the utility losses fall below some threshold. Hence, the lack of actual bunching does not necessarily imply that there is no bunching of desired taxable income.

If the first interpretation is correct (lower compensated elasticity), marginal deadweight 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 now elaborate on this second point.

4.2. 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. This 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 piecewise 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 is 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. The 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 cancelled vacation plans, better health than expected, assigned overtime, an unexpectedly large Christmas bonus, etc.

Let us write the realized taxable income as $A^r = A^d + \varepsilon$, where A^d is the desired taxable income and ε is the shock or optimization error. Let us make the reasonable assumption that the shocks are independent of desired hours and the tax system. Then, $\partial A^r / \partial \gamma = \partial A^d / \partial \gamma$ for any tax parameter γ 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 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 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 therefore 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, in Section 4.6, we perform computations under the assumption that each individual plans desired taxable income and then faces randomness. In an Appendix, we describe the decision problem of an individual taking account of the optimization errors when deciding on desired taxable income. We leave it to future research to explore whether the two different ways to take account of the optimization errors matter for the marginal deadweight loss calculations.

4.3. The taxable income elasticity

Deadweight loss measures depend on the elasticity of taxable income. However, there is no consensus about the size of the latter. Feldstein (1999) estimated the taxable income elasticity to 3. This is by most scholars considered to be much too high. Recent studies report much lower elasticities. Using Swedish data and (commonly used) strong functional form assumptions, Blomquist and Selin (2010) find a compensated elasticity for married males of 0.24. Using non-parametric methods and fully allowing for individual heterogeneity Blomquist et al. (2015) estimate an average compensated elasticity of 0.53. Using German data, Doerrenberg et al. (2015) report an elasticity between 0.54–0.68. For the US, Weber (2014) finds an elasticity of 0.86, Kumar and Liang (2015) an elasticity of 0.75 and Burns and Ziliak (2017) an elasticity of 0.55. Given the large spread in estimated elasticities, we have chosen to report results for four values of the taxable income elasticity, ranging from 0.2 to 0.8. The reader is welcomed to focus on the value he or she regards as most reasonable. For the US, we believe that the results for an elasticity of 0.6–0.8 are of most interest. In the tables below, we show the bias in various measures of the deadweight loss.

4.4. Aggregating the marginal deadweight loss for the population

Due to heterogeneity in preferences, individuals facing the same tax schedule chose different taxable incomes, which generates different marginal deadweight losses. From a welfare point of view, there is no obvious way how one should aggregate the marginal deadweight loss for different individuals (see Hausman and Newey, 2016). We do not contribute to this important research question, but instead follow a fairly common procedure and calculate the average of the marginal deadweight losses of the different agents in the population. Whatever the weights that are used, it is clear that the aggregate marginal deadweight loss obtained with the function A_L^h gives a higher value than if calculated using A^h .

As explained in Section 3, there is no clear way of aggregating the marginal deadweight loss per marginal tax dollar. We below compute the marginal deadweight loss per tax dollar as the ratio between the average marginal deadweight loss and the average marginal tax revenue.

4.5. Parametrization of the tax system, individual preferences and skill levels

We are using the Internet TaxSim simulator, provided by NBER, to recover the budget constraint for the US. This allowed us to 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. More precisely, we used the Californian tax schedule to compute state taxes. 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 1993 and 2006.⁷ The budget set exhibits nonconvexities in 1993 and 2006. As noted before, on these income intervals, the standard linearization procedure would underestimate the correct marginal deadweight loss. These nonconvexities are sharp in 1993 and we therefore fully took them into account in our computations. By contrast, they are small in 2006 and it was therefore innocuous to convexify the budget set by taking its convex hull. We obtain budget constraints with 18, 20 and 13 kinks, in 1979, 1994 and 2006 respectively.

Regarding individual preferences, we follow Saez (2010) and consider that individual preferences are described by a quasilinear and isoelastic function of the form:

$$U(C, A, v) = C - \frac{v}{1 + 1/\beta} \left(\frac{A}{v}\right)^{1+1/\beta} \quad (19)$$

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 . This utility function is by far the most used utility function in the taxable income literature and convenient in numerical calculations as it implies closed form solutions for the compensated taxable income function and the derivatives needed for the calculations of the marginal deadweight loss. As discussed above, there is no consensus in the empirical literature about the size of the taxable income elasticity. We provide computations for different plausible values of β . Given (19), the marginal deadweight loss of a v -agent with optimum along a linear segment of slope θ is equal to: $(1 - \theta) \beta v \theta^{\beta-1}$; that of an agent with desired taxable income located at a kink is 0.

We use data from the CPS labor extracts 1979, 1993 and 2006 to calibrate the wage parameter v . Following Mirrlees (1971), we assume that its distribution is lognormal. What we observe in the data is realized incomes A^r . Following Burtless and Hausman (1978), we assume that, for a given person, desired income differs from realized income by an amount ε , with $A^r = A^d + \varepsilon$. Using (19), we obtain $A^d = v \theta^\beta$ along a linear segment of the budget constraint with slope θ . The underlying v is therefore given by $v = (A^r - \varepsilon) / \theta^\beta$. We further assume: $\varepsilon \sim N(\mu, \sigma)$. If we denote by $\tilde{\theta}$ the slope of the budget constraint faced by the individual with average realized gross earnings $E(A^r)$, we obtain $E(v) = E(A^r) / \tilde{\theta}^\beta$ and $\text{Var}(v) = (\text{Var}(A^r) - \sigma^2) / \tilde{\theta}^{2\beta}$. The standard deviation σ of the error term ε is an important parameter. There is however no

⁷ The payroll tax and the sales taxes are linear (however, above an annually adjusted threshold, the payroll tax only consists of the Medicare tax of 2.9%). The payroll tax (FICA) is 15.3% in 1993 and 2006, and 12.26% in 1979. The Californian sales tax is 7.25%. Local taxes vary. In our computations, we assume that the local sales tax is 0.5%.

Table 1
Results for 1979.

β	0.2	0.4	0.6	0.8
% people kinks	6.88%	13.24%	19.16%	24.43%
MDW correct	2887	5723	8493	11195
MDW linear	3072	6474	10179	14162
Bias	6.42%	13.11%	19.85%	26.50%
MTR correct	13185	10931	8689	6467
MTR linear	13000	10181	7003	3500
Bias	-1.41%	-6.87%	-19.40%	-45.88%
MDW/\$ correct	0.219	0.534	0.977	1.73
MDW/\$ linear	0.236	0.636	1.454	4.05
Bias	7.93%	21.45%	48.71%	133.73%

Table 2
Results for 1993.

β	0.2	0.4	0.6	0.8
% people at kinks	5.31%	9.79%	11.16%	14.66%
MDW correct	4645	7303	6943	9729
MDW linear	4943	8232	8089	11999
Bias	6.41%	12.72%	16.51%	23.33%
MTR correct	19058	13680	8526	7015
MTR linear	18760	12752	7380	4745
Bias	-1.56%	-6.78%	-13.44%	-32.36%
MDW/\$ correct	0.244	0.534	0.814	1.387
MDW/\$ linear	0.263	0.645	1.096	2.529
Bias	8.10%	20.92%	34.60%	82.33%

Table 3
Results for 2006.

β	0.2	0.4	0.6	0.8
% people at kinks	3.66%	7.07%	9.93%	12.66%
MDW correct	7344	13858	17125	23282
MDW linear	7681	15052	19277	27287
Bias	4.59%	8.62%	12.56%	17.20%
MTR correct	26842	19985	13249	8585
MTR linear	26505	18791	11097	4580
Bias	-1.25%	-5.98%	-16.24%	-46.65%
MDW/\$ correct	0.274	0.693	1.293	2.71
MDW/\$ linear	0.290	0.801	1.737	5.96
Bias	5.91%	15.52%	34.39%	119.72%

obvious way to calibrate it. We assume that it is such that 10% of the agents face a positive shock larger than 10% of average gross income. We get $\sigma = 859$ for 1979, $\sigma = 1227$ for 1993 and $\sigma = 2787$ for 2006. We below provide a sensitivity analysis for the 2006 tax system. Note that we calibrate the distribution of wages for each year and each value of β . If we used for a given year the same skill distribution for the different values of β , the distribution of desired taxable income would be located on different parts of the tax schedule and results would partly depend on that fact.

The numerical results below implement the formulas derived in Section 3.2. Our calculations for the linearization procedure build on the assumption that optimization errors are the reason for no observations at the kink points. When calculating the marginal deadweight loss using the linearization technique we therefore add optimization errors to the taxable incomes at the kink points. Assuming the optimization errors to be normally distributed with mean zero this implies that 50% of the kink point solutions will be allocated to the linear segment to the left of the kink and 50% to the linear segment to the right.

4.6. Results

In Tables 1–3, we present the results for the three years. For all years, looking at the rows showing the correct marginal deadweight loss per marginal tax revenue dollar, we see how it increases with the taxable income elasticity. For example, for 2006, the marginal deadweight loss per marginal tax revenue dollar goes from 27.4 cent for $\beta = 0.2$ to 2.710 dollars for $\beta = 0.8$. This underscores the importance of having correct information about the taxable income elasticity. For the rows showing the numbers where the linearization procedure is used, the increase in the marginal deadweight loss per marginal tax revenue dollar is even larger. For 2006 it goes from 29 cents for $\beta = 0.2$ to 5.91 dollars for $\beta = 0.8$. Looking at the biases the linearization procedure gives rise to, we see that the biases for the marginal deadweight loss are small (negligible) for β equal to 0.2 or 0.4. The same is true for the biases for the marginal tax revenue. However, for β equal to 0.6

Table 4
Sensitivity analysis for 2006 and $\beta = 0.6$.

σ	500 (–82%)	$\sigma_b = 2787$ (benchmark)	4000 (+43%)	6000 (+115%)
% people at kinks	9.38%	9.93%	9.28%	9.46%
MDW correct	17159	17125	17074	17006
MDW linear	19306	19277	19190	19179
Bias	12.51%	12.56%	12.40%	12.77%
MTR correct	13241	13249	13245	13350
MTR linear	11095	11097	11130	11177
Bias	–16.21%	–16.24%	–15.97%	–16.28%
MDW/\$ correct	1.296	1.293	1.289	1.274
MDW/\$ linear	1.740	1.737	1.724	1.716
Bias	34.28%	34.39%	33.75%	34.70%

and 0.8 the biases become more serious. For example, for 1979 and $\beta = 0.8$, the linearization procedure overestimates the marginal deadweight loss by almost 27% and underestimates the marginal tax revenue by –46%. The real problem with the linearization procedure arises for the computations of the marginal deadweight loss per marginal tax revenue dollar. There, the biases for the marginal deadweight loss and the marginal tax revenue compound and in 1979 give rise to an overestimate of 134% for the marginal deadweight loss per marginal tax revenue dollar if $\beta = 0.8$. For 2006 the overestimate is 120%. Comparing the marginal deadweight losses across the three tax systems, we see that it is the 1993 tax system that gives rise to the lowest marginal deadweight loss per marginal tax dollar and the 2006 tax system the highest. For $\beta = 0.8$, the marginal deadweight loss per marginal tax dollar is 1.39 dollars for the 1993 tax system and 2.71 dollars for the 2006 tax system.

In Table 4, we examine the sensitivity of our calibration with respect to changes in the standard deviation of the optimization error for an elasticity of 0.6 and for year 2006. We consider one large deviation to the left ($\sigma = 500$) and two large deviations to the right ($\sigma = 4000$ and 6000), compared to our benchmark scenario where $\sigma = 2787$. The results of our benchmark calibration are quite robust to relatively large variations in the standard deviation of the optimization error. For example, for an increase in σ by 115% from the benchmark value of 2787 to 6000, the correct marginal deadweight loss decreases by 0.69% and the correct marginal tax revenue increases by 0.76%.

5. Concluding comments

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 therefore of interest to know how 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. In this article, we first derive the correct way to calculate the marginal deadweight loss when the budget constraint is smooth and convex and we do this for a change in any tax parameter. 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.

A common procedure to calculate the marginal deadweight loss of a change in a tax parameter has been to linearize the budget constraint and then calculate the marginal deadweight loss for a variation in the linearized budget constraint. We show that such a procedure does not give the correct value of the marginal deadweight loss and we derive formulas that describe the resulting biases.

We next describe how to calculate the marginal deadweight loss for a piecewise linear tax. We do this for a particular type of tax change, 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. We show that for those with desired taxable income in the interior of a linear segment the marginal deadweight loss can be calculated in the same way as for a linear budget constraint. For those with desired taxable income at a kink point the marginal deadweight loss will be zero. The effect of the change in the tax parameter will be just as if it were a lump-sum tax. 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. When calculating the average marginal deadweight loss for a population, it is therefore crucial to find out how many individuals there are with desired hours at a kink point. Earlier empirical studies of the marginal deadweight loss have neglected the issue of optimization errors and implicitly assumed that all actual taxable incomes are the same as desired taxable income, implying that virtually all observations have been treated as if desired taxable income is in the interior of a linear segment.

We finally perform numerical computations for the US tax system in 1979, 1993 and 2006. We find that for a taxable income elasticity of 0.6 the bias in marginal deadweight loss per marginal tax dollar from linearizing the budget constraint

can be up to 50% and if the taxable income elasticity is 0.8 it can be well over 100%. Because of the potentially misleading policy implications of the linearization procedure, we believe that the curvature of the tax system should be fully accounted for when measuring the deadweight loss. It is easy 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.

Acknowledgments

We are grateful to two anonymous referees, the editor Whitney Newey, Spencer Bastani, Thomas Gaube, Aart Gerritsen, Laurence Jacquet, Henrik Kleven, Kai Konrad, Klaus Thustrup Kreiner, Jean-Marie Lozachmeur, Olof Johansson Stenman, Etienne Lehmann, Luca Micheletto, Håkan Selin and Alain Trannoy for very helpful comments and suggestions.

Appendix. Anticipated random shocks

As explained in the text, realized income may differ from planned income for various reasons. In this [Appendix](#), we consider that individuals face random shocks, know the distribution of them, and behave as expected utility maximizers. Following [Saez \(1999\)](#), we consider that the random shocks correspond to increases or decreases in planned labor earnings. For simplicity we assume these shocks do not affect the disutility of labor. An unexpected high Christmas bonus would be an example of positive shock and an unexpected low Christmas bonus would be an example of a negative shock. We in principle could also consider shocks which both modify effort and earnings.

Let the shocks ε have support $(\underline{\varepsilon}, \bar{\varepsilon})$, probability density function $e(\varepsilon)$, and expected value 0. Planned or desired income A^d is therefore the solution to:

$$\max_A \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[A + \varepsilon - T(A + \varepsilon) - \frac{v}{1 + \frac{1}{\beta}} \left(\frac{A}{v} \right)^{1 + \frac{1}{\beta}} \right] e(\varepsilon) d\varepsilon. \quad (\text{A.1})$$

We define the expected tax liability as:

$$\hat{T}(A) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} T(A + \varepsilon) e(\varepsilon) d\varepsilon$$

and the expected marginal tax rate $\hat{T}'(A)$ as $d\hat{T}(A)/dA$. Using the first-order condition for [\(A.1\)](#), we obtain desired income $A^d = v \times \left(1 - \hat{T}'(A^d) \right)^{\beta}$. This implies that, in order to make their choice, individuals do not consider the actual tax schedule, but the expected smooth one, in which actual marginal tax rates are replaced by $\hat{T}'(A^d)$.

References

- Auerbach, A., Hines, J., 2002. Taxation and economic efficiency. In: Auerbach, A.J., Feldstein, M. (Eds.), *Handbook of Public Economics*, Vol. 3. Elsevier, 1347–1421.
- Bastani, S., Selin, H., 2014. Bunching and non-bunching at kink points of the Swedish tax schedule. *J. Publ. Econ.* 109, 36–49.
- Blomquist, S., 1983. The effect of income taxation on the labor supply of married men in Sweden. *J. Publ. Econ.* 22, 169–197.
- Blomquist, S., Kumar, A., Liang, C.-Y., Newey, W.K., 2015. Individual Heterogeneity, Nonlinear Budget Sets, and Taxable Income. CESifo Working Paper No. 5320.
- Blomquist, S., Newey, W., 2002. Nonparametric estimation with nonlinear budget sets. *Econometrica* 70, 2455–2480.
- Blomquist, S., Selin, H., 2010. Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. *J. Publ. Econ.* 94, 878–889.
- Blundell, R., Hoynes, H., 2004. Has 'in-work' benefit reform helped the labor market? In: Card, D., Blundell, R., Freeman, R.B. (Eds.), *Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980–2000*. University of Chicago Press, pp. 411–459.
- Burns, S.K., Ziliak, J.P., 2017. Identifying the elasticity of taxable income. *Econom. J.* 127, 297–329.
- Burtless, G., Hausman, J.A., 1978. The effect of taxation on labor supply: evaluating the Gary negative income tax experiment. *J. Polit. Econ.* 86, 1103–1130.
- Burtless, G., Moffitt, R.A., 1984. The effect of social security benefits on labor supply of the aged. In: Aaron, H., Burtless, G. (Eds.), *Retirement and Economic Behavior*. The Brookings Institution, Washington, DC.
- Chapman, B., Leigh, A., 2009. Do very high tax rates induce bunching? Implications for the design of the contingent loan schemes. *Econ. Rec.* 85, 276–289.
- Chetty, R., 2012. Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. *Econometrica* 80, 969–1018.
- Dahlby, B., 1998. Progressive taxation and the social marginal cost of public funds. *J. Publ. Econ.* 67, 105–122.
- Dahlby, B., 2008. *The Marginal Cost of Public Funds, Theory and Applications*. MIT Press.
- Doerrenberg, P., Peichl, A., Siegloch, S., 2015. The elasticity of taxable income in the presence of deduction possibilities. *J. Publ. Econ.* 151, 41–55.
- Dupuit, J., 1844. De la mesure de l'utilité des travaux publics. 2, 332–375.
- Eissa, N., Kleven, H., Kreiner, C., 2008. Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers. *J. Publ. Econ.* 92, 795–816.
- Feldstein, M., 1995. Behavioral responses to tax rates: evidence from TRA86. *Amer. Econ. Rev.* 85, 170–174.
- Feldstein, M., 1999. Tax avoidance and the deadweight loss of the income tax. *Rev. Econ. Stat.* 81, 674–680.
- Friedberg, L., 2000. The labor supply effects of the social security earnings test. *Rev. Econ. Stat.* 82, 48–63.
- Gelber, A., Mitchell, J.W., 2012. Taxes and time allocation: evidence from single women and men. *Rev. Econom. Stud.* 79, 863–897.
- Gruber, J., Saez, E., 2002. The elasticity of taxable income: evidence and implications. *J. Publ. Econ.* 84, 1–32.
- Harberger, A., 1962. The incidence of the corporation income tax. *J. Polit. Econ.* 70, 215–240.
- Harberger, A., 1964. The measurement of waste. *Amer. Econ. Rev.* 54, 58–76.
- Hausman, J.A., 1979. The econometrics of labor supply on convex budget sets. *Econom. Lett.* 3, 171–174.
- Hausman, J.A., 1985. The econometrics of nonlinear budget sets. *Econometrica* 53, 1255–1282.

- Hausman, J.A., Newey, W., 1995. Nonparametric estimation of exact consumer surplus and deadweight loss. *Econometrica* 63, 1445–1476.
- Hausman, J.A., Newey, W., 2016. Individual heterogeneity and average welfare. *Econometrica* 84, 1225–1248.
- Kleven, H., Kreiner, C., 2006. The marginal cost of public funds: hours of work versus labor force participation. *J. Publ. Econ.* 90, 1955–1973.
- Kopczuk, W., 2005. Tax bases, tax rates and the elasticity of reported income. *J. Publ. Econ.* 89, 2093–2119.
- Kumar, A., Liang, C.-Y., 2015. The Taxable Income Elasticity: A Structural Differencing Approach. UCFS Working Paper 2015:1.
- Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation. *Rev. Econom. Stud.* 38, 175–208.
- Saez, E., 1999. Do Taxpayers Bunch at Kink Points? NBER Working Paper 7366 Working Paper 7366.
- Saez, E., 2010. Do taxpayers bunch at kink points? *Am. Econ. J. Econ. Policy* 2, 180–212.
- Saez, E., Slemrod, J., Giertz, S., 2012. The elasticity of taxable income with respect to marginal tax rates: a critical review. *J. Econ. Lit.* 50, 3–50.
- Weber, C.E., 2014. Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. *J. Publ. Econ.* 117, 90–103.