UNIVERSITY *of York*

Discussion Papers in Economics

No. 18/02

Locally Optimal Three-Bracket Piecewise Linear Income Taxation

Alan Krause

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Locally Optimal Three-Bracket Piecewise Linear Income Taxation*

Alan Krause†
University of York

13 February 2018

Abstract

The aim of this paper is to examine the setting of income tax policy from the perspective faced by governments. The government takes the current income tax schedule as the starting point, and seeks to implement a small change in the tax schedule that is both feasible and desirable. If no such change is possible, the current income tax schedule is said to be locally optimal, because it cannot be improved upon via a small reform. We assume that the current income tax schedule is piecewise linear with three tax brackets, which approximates most real-world income tax schedules. The characteristics of locally-optimal piecewise linear income tax schedules are then derived, with particular attention paid to the extent to which they depart from linearity.

Keywords: piecewise linear income taxation; tax reform.

JEL Classifications: H21, H24.

*This paper is a substantial revision of previous drafts under the titles "Nonlinear Income Tax Reforms" and "Piecewise Linear Income Tax Reforms". The current paper replaces those previous drafts. I thank Peter Simmons, Jacco Thijssen, Zaifu Yang, and seminar participants at the University of York for their helpful comments. Any errors remain my responsibility.

†Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K. E-mail: alan.krause@york.ac.uk.

1 Introduction

The aim of this paper is to examine the setting of income tax policy from the perspective faced by governments. This perspective can be summarised as follows. Given the current income tax system, is it possible to implement a small change in the tax schedule that is both feasible and desirable? To address this question, we assume that the current income tax system is a progressive piecewise linear income tax schedule with three tax brackets.[1] Such a system closely approximates the vast majority of real-world income tax systems. We then examine if there exists a small (modelled as differential) change in the tax system that is both feasible (maintains budget balance) and desirable (welfare improving). If such change is not possible, the current tax system is said to be locally optimal because it cannot be improved upon, at least not via a small reform. The characteristics of locally-optimal piecewise linear income tax schedules are then discussed, with particular attention paid to the extent to which they depart from linearity.

Our approach differs from that taken in the optimal tax literature. Under optimal tax analyses, the current tax system imposes no constraints and the new tax system is effectively designed from scratch. Indeed, the Mirrlees (1971) approach to optimal income taxation imposes no constraints on the government other than that it cannot observe each individual's skill type (thus preventing the implementation of personalised lump-sum taxes and transfers as the Second Welfare Theorem would recommend). However, it is more realistic to assume that the government starts with the current tax system, and is constrained to make small changes in taxes. The tax reform literature, beginning with Guesnerie (1977, 1995),[2] seeks to incorporate such behavioural constraints on the government. Our analysis is an exercise in tax reform.

Compared to the extensive literature on optimal nonlinear income taxation using the Mirrlees model, the literature on optimal piecewise linear income taxation is scarce.

[1] Throughout the paper, we use the terms 'progressive' and 'regressive' as in common parlance. That is, a progressive income tax schedule is one with marginal tax rates increasing in income. A regressive income tax schedule has declining marginal tax rates.

[2] There are a number of other papers that use tax reform methods to examine various policy questions. See e.g. Feldstein (1976), Hatta (1977), Diewert (1978), Weymark (1979), Konishi (1995), Brett (1998), Blackorby and Brett (2000), Murty and Russell (2005), Fleurbaey (2006), Krause (2007, 2009), and Duclos, et al. (2008). For a textbook treatment of tax reform analysis, see chapter 6 in Myles (1995).

Sheshinski (1989) derives the optimal two-bracket piecewise linear income tax schedule, and claims that it is necessarily convex (progressive). That is, the marginal tax rate applicable in the high-income tax bracket must be greater than that applicable in the low-income tax bracket. Slemrod, et al. (1994) claim that Sheshinski's analysis is flawed; they reach the opposite conclusion that the optimal two-bracket piecewise linear income tax schedule is non-convex (regressive). More recently, Apps, et al. (2014) show that optimal two-bracket piecewise linear income taxation may be convex or non-convex, but under an empirically-plausible parameterisation it is convex. Our analysis differs from those papers in the following respects. First, we consider three tax brackets rather than two.[3] Most real-world income tax systems seek to tax low-income, middle-income, and high-income workers at different rates, and our model captures that key feature. Second, as mentioned above, our analysis uses tax reform methods. We assume that the current income tax system is piecewise linear, has three brackets, and is convex (progressive), which describes most real-world income tax systems. We then derive the conditions under which there does not exist a small change in the tax schedule that is both feasible and desirable, i.e., the conditions under which the current income tax system is locally optimal.

Our main result can be summarised as follows. The extent to which a locally-optimal piecewise linear income tax schedule departs from linearity depends upon the populations of 'special' individuals. These are individuals who choose to earn income at the kink points of the tax schedule. These individuals appear in some of the tax formulas in the related literature,[4] but their importance has not been highlighted. Instead, the focus of the previous literature is on determining whether convex or non-convex income tax schedules are globally optimal. In our model, the current income tax schedule is necessarily convex (progressive). The individuals who choose to earn at the kink points can then be interpreted as individuals who would like to earn a little more income, but are deterred from doing so by the higher marginal tax rate that they would face. Our

[3] Dahlby (1998) considers progressive piecewise linear income taxation with n tax brackets, but her focus is on the marginal cost of public funds rather than the optimality (or otherwise) of the income tax schedule.

[4] See, for example, equation (27) in Apps, et al. (2014).

analytical approach clearly demonstrates the importance of these individuals: the larger their populations, the greater a locally-optimal piecewise linear income tax schedule departs from linearity. This raises an interesting issue. Suppose the current income tax schedule is approximately linear. Then one would expect that relatively few individuals will bunch at the kink points, because the increase in the marginal tax rate from one tax bracket to the next is small. This case, with few individuals earning at the kink points and an approximately linear income tax schedule, is consistent with local optimality. But likewise, suppose the current income tax schedule is far from linear, i.e., there are large changes in marginal tax rates from one tax bracket to the next. Then one would expect that many individuals will bunch at the kink points. That situation is also consistent with local optimality. Accordingly, current piecewise linear income tax schedules, whether approximately linear or not, can be locally optimal.

The remainder of the paper is organised as follows. Section 2 presents our analytical framework, Section 3 derives our results, and Section 4 presents two numerical examples. Section 5 discusses some problems that would arise if using our methodology to analyse regressive or U-shaped income tax schedules. Section 6 concludes, while proofs and some other mathematical details are provided in an appendix.

2 Analytical Framework

There are five types of individual who are distinguished by their wages, $w_5 > w_4 > w_3 > w_2 > w_1 > 0$. Let n_i denote the number of type i individuals. The size of the population is normalised to unity, so that $\sum_{i=1}^{5} n_i = 1$. Type i's income is denoted by y_i, where $y_i = w_i l_i$ with l_i representing type i's labour supply. Suppose the current income tax system is a progressive piecewise linear schedule with three tax brackets. Figure 1 illustrates this case, in which income up to and including $\widehat{y}$ is taxed at a marginal rate t_1, income greater than $\widehat{y}$ but less than or equal to $\overline{y}$ is taxed at a marginal rate t_2, and all income above $\overline{y}$ is taxed at a marginal rate t_3. As $t_3 > t_2 > t_1$, progressive taxation generates a convex budget set (thick dashed line), where x_i denotes type i's post-tax income (consumption). Type 1 individuals represent low-income workers who choose to earn less than $\widehat{y}$. Their optimal choice y_1 is determined by the usual tangency condition between their indifference curve and budget line. Likewise, type 3 individuals represent middle-income workers who choose to earn y_3, which is between $\widehat{y}$ and $\overline{y}$. Type 5 individuals represent high-income workers who choose to earn $y_5 > \overline{y}$. Type 2 and type 4 individuals are 'special' types. These types are included in our model to represent individuals who choose to earn at the kink points, $y_2 = \widehat{y}$ and $y_4 = \overline{y}$. As in Apps, et al. (2014), type 2 individuals can be interpreted as representing those workers who would like to earn a little more than $\widehat{y}$, but are deterred from doing so by the higher marginal tax rate t_2 that they would face. These workers are effectively constrained to earn $\widehat{y}$. Analogously, type 4 workers are deterred by the higher marginal tax rate t_3, and are effectively constrained to earn $\overline{y}$.

All individuals have the same preferences, represented by the quasi-linear utility function:

$$x_i - v(l_i) \tag{2.1}$$

where $v(\cdot)$ is increasing and strictly convex. Under three-bracket piecewise linear income taxation, the budget constraints faced by the individuals can be written as:

$$x_i \leq a + (1 - t_1)y_i \qquad for \qquad y_i \leq \widehat{y} \tag{2.2}$$

$$x_i \leq a + (1 - t_1)\widehat{y} + (1 - t_2)(y_i - \widehat{y}) \qquad for \qquad \widehat{y} < y_i \leq \overline{y} \tag{2.3}$$

$$x_i \leq a + (1 - t_1)\widehat{y} + (1 - t_2)(\overline{y} - \widehat{y}) + (1 - t_3)(y_i - \overline{y}) \qquad for \qquad y_i > \overline{y} \tag{2.4}$$

where a is a lump-sum tax/transfer received by each individual.

3 Locally-Optimal Income Tax Schedules

Suppose the current piecewise linear income tax system is progressive ($t_3 > t_2 > t_1 > 0$), as illustrated in Figure 1. As type 1 individuals have chosen to earn y_1, they have acted as if maximising the utility function (2.1) subject to the budget constraint (2.2). The solution yields the function $y_1(w_1, t_1)$. Likewise, type 3 individuals have acted as if maximising (2.1) subject to (2.3), yielding $y_3(w_3, t_2)$, and type 5 individuals have acted as if maximising (2.1) subject to (2.4), yielding $y_5(w_5, t_3)$. As discussed above, type 2 and type 4 individuals can be interpreted as being constrained to earn $\widehat{y}$ and $\overline{y}$, respectively. Therefore, type 2 individuals have acted as if maximising (2.1) subject to (2.2) and the constraint that $y_2 \leq \widehat{y}$, while type 4 individuals have acted as if maximising (2.1) subject to (2.3) and the constraint that $y_4 \leq \overline{y}$.[5]

The current tax system is feasible if tax revenues are greater than or equal to expenditures:

$$n_1 t_1 y_1(w_1, t_1) + n_2 t_1 \widehat{y} + n_3 \left[(t_1 - t_2)\widehat{y} + t_2 y_3(w_3, t_2)\right] + n_4 \left[(t_1 - t_2)\widehat{y} + t_2 \overline{y}\right]$$

$$+ n_5 \left[(t_1 - t_2)\widehat{y} + (t_2 - t_3)\overline{y} + t_3 y_5(w_5, t_3)\right] - a - g \geq 0 \tag{3.1}$$

The first five terms in equation (3.1) are total tax payments by each type of individual, a is total lump-sum payments by the government (recall that the size of the population is normalised to one), and g is the government's exogenous revenue requirement. We assume that equation (3.1) holds with equality, i.e., the government's budget is exactly balanced.

The vector $\mathrm{d}R := \langle \mathrm{d}t_1, \mathrm{d}t_2, \mathrm{d}t_3, \mathrm{d}\widehat{y}, \mathrm{d}\overline{y}, \mathrm{d}a \rangle$ is a tax reform, which represents a small (differential) change in the income tax schedule. A tax reform is feasible if and only if:

$$\nabla Z \mathrm{d}R \geq 0 \tag{3.2}$$

where ∇Z is the gradient of equation (3.1) and is:

$$\nabla Z := \langle n_1 \left(y_1 + t_1 \frac{\partial y_1}{\partial t_1} \right) + (1 - n_1)\widehat{y},\ n_3 \left(y_3 + t_2 \frac{\partial y_3}{\partial t_2} \right) - (n_3 + n_4 + n_5)\widehat{y} + (n_4 + n_5)\overline{y},$$

$$n_5 \left(y_5 + t_3 \frac{\partial y_5}{\partial t_3} \right) - n_5 \overline{y},\ n_2 t_1 + (n_3 + n_4 + n_5)(t_1 - t_2),\ n_4 t_2 + n_5 (t_2 - t_3),\ -1 \rangle \tag{3.3}$$

Therefore, a tax reform is feasible if it does not move the government's budget position

[5] The details of these maximisation problems are provided in the appendix.

into deficit.

Suppose social welfare is measurable by a weighted utilitarian social welfare function, $W = \sum_{i=1}^{5} \pi_i n_i V_i$, where π_i is type i's welfare weight (with $\pi_1 > \pi_2 > \pi_3 > \pi_4 > \pi_5 > 0$) and V_i is type i's indirect utility function.[6] A tax reform is desirable if and only if:

$$\nabla W \mathrm{d}R > 0 \tag{3.4}$$

where:

$$\begin{aligned}\nabla W := \langle &-\pi_1 n_1 y_1 - \widehat{y} \sum_{i=2}^{5} \pi_i n_i,\ \pi_3 n_3(\widehat{y} - y_3) + (\pi_4 n_4 + \pi_5 n_5)(\widehat{y} - \overline{y}),\ \pi_5 n_5(\overline{y} - y_5), \\ &\pi_2 n_2 \left[1 - t_1 - v'(\frac{\widehat{y}}{w_2})\frac{1}{w_2}\right] + (t_2 - t_1) \sum_{i=3}^{5} \pi_i n_i,\ \pi_4 n_4 \left[1 - t_2 - v'(\frac{\overline{y}}{w_4})\frac{1}{w_4}\right] + \pi_5 n_5 (t_3 - t_2),\ \sum_{i=1}^{5} \pi_i n_i \rangle\end{aligned} \tag{3.5}$$

is the gradient of the social welfare function.[7] Therefore, a tax reform is desirable if it increases social welfare.

If there does not exist a tax reform $\mathrm{d}R$ that satisfies (3.2) and (3.4), then there does not exist a reform that is both feasible and desirable. The current piecewise linear income tax schedule is then said to be locally optimal. By Motzkin's Theorem of the Alternative,[8] if there does not exist a tax reform $\mathrm{d}R$ that satisfies (3.2) and (3.4), then there exists $\theta \geq 0$ and $\beta > 0$ such that:

$$\theta \nabla Z + \beta \nabla W = 0^{(6)} \tag{3.6}$$

Using equation (3.6) it is shown in the appendix that:

Proposition 1 *Suppose the current income tax schedule is piecewise linear, has three tax brackets, and is progressive* ($t_3 > t_2 > t_1 > 0$). *Further suppose that social welfare is measurable by a weighted utilitarian social welfare function. Then the current income*

[6] As in much of the related literature, we assume for analytical convenience that the direct utility function is quasi-linear in consumption. However, as is well known, this precludes the use of a pure utilitarian social welfare function. We therefore use a weighted utilitarian social welfare function.

[7] The details underlying the derivation of equation (3.5) are provided in the appendix.

[8] A statement of Motzkin's Theorem is provided in the appendix.

tax schedule: (i) *may be locally optimal, and* (ii) *if it is locally optimal*:

$$t_2 - t_1 = \frac{n_2\left(\pi_2\left[1 - t_1 - v'(\frac{\hat{y}}{w_2})\frac{1}{w_2}\right] + t_1\sum_{i=1}^{5}\pi_i n_i\right)}{\sum_{i=1}^{2}\pi_i n_i - (n_1+n_2)\sum_{i=1}^{5}\pi_i n_i} \tag{3.7}$$

$$t_3 - t_2 = \frac{n_4\left(\pi_4\left[1 - t_2 - v'(\frac{\overline{y}}{w_4})\frac{1}{w_4}\right] + t_2\sum_{i=1}^{5}\pi_i n_i\right)}{n_5\sum_{i=1}^{4} n_i(\pi_i - \pi_5)} \tag{3.8}$$

Part (i) of Proposition 1 establishes that a convex (progressive) piecewise linear income tax schedule can be locally optimal.[9] Part (ii) of Proposition 1 provides equations that show the extent to which a locally-optimal piecewise linear income tax schedule departs from linearity, i.e., expressions for $t_2 - t_1$ and $t_3 - t_2$. While these expressions are a little complex, it can be seen that the populations of the 'special' types, n_2 and n_4, appear to play an important role in determining the departure from linearity. This can be seen very clearly if the social welfare function is maxi-min:

Proposition 2 *Suppose the current income tax schedule is piecewise linear, has three tax brackets, and is progressive* ($t_3 > t_2 > t_1 > 0$). *Further suppose that social welfare is measurable by a maxi-min social welfare function. Then the current income tax schedule*: (i) *may be locally optimal, and* (ii) *if it is locally optimal*:

$$\frac{t_2 - t_1}{t_1} = \frac{n_2}{n_3 + n_4 + n_5} \qquad \textit{and} \qquad \frac{t_3 - t_2}{t_2} = \frac{n_4}{n_5} \tag{3.9}$$

$$n_4 \geq n_2 \qquad \Longrightarrow \qquad \frac{t_3 - t_2}{t_2} > \frac{t_2 - t_1}{t_1} \tag{3.10}$$

In sum, the larger the populations of type 2 and type 4 individuals, the greater the extent to which locally-optimal piecewise linear income taxation departs from linearity. To see the intuition, consider for example the maxi-min case, but suppose that:

$$\frac{t_2 - t_1}{t_1} < \frac{n_2}{n_3 + n_4 + n_5} \qquad \Longrightarrow \qquad t_1 n_2 > (t_2 - t_1)(n_3 + n_4 + n_5) \tag{3.11}$$

[9] While this might seem obvious, it is worth recalling that some of the previous literature (e.g., Slemrod, et al. 1994) has suggested that convex piecewise linear income taxation is not optimal. Moreover, in Mirrlees-style nonlinear income tax models, it is often found that a declining (e.g., Mankiw, et al. 2009) or U-shaped (e.g., Diamond 1998 and Saez 2001) pattern of marginal tax rates is optimal.

The situation depicted in equation (3.11) is not consistent with local optimality. The government could implement a small increase in the first tax bracket ($\mathrm{d}\widehat{y} > 0$), enabling type 2 individuals to work longer and thus increasing tax revenues by $t_1 n_2$. This increase would exceed the loss of tax revenues collected from type 3, 4, and 5 individuals, the loss being $(t_2 - t_1)(n_3 + n_4 + n_5)$, who now have more of their incomes taxed at the marginal rate t_1. Local optimality requires that the benefits of any possible change be equal to the costs, and this condition creates a relationship between differences in the marginal tax rates and the populations of type 2 and type 4 individuals. A similar intuition applies under a weighted utilitarian social welfare function, except in that case the benefits and costs include welfare effects.

Under maxi-min, one can derive a simple expression for the degree of progressivity of the income tax schedule at a local optimum. As equation (3.10) shows, $n_4 \geq n_2$ is sufficient (not necessary) for the degree of progressivity to be increasing, as measured by the difference between $(t_3 - t_2)/t_2$ and $(t_2 - t_1)/t_1$. However, this basic condition does not hold under a weighted utilitarian social welfare objective, as demonstrated by the numerical examples below in Section 4.

Propositions 1 and 2 lend support to the view that current real-world income tax schedules, whether approximately linear or not, can be locally optimal. If the current income tax schedule is approximately linear, then one would expect little bunching at the kink points because the increase in the marginal tax rate from one tax bracket to the next is small. This situation is consistent with local optimality. But likewise, if the current income tax schedule is far from linear, with large kink points, then one would expect significant bunching. That situation is also consistent with local optimality.

There is some empirical evidence that supports the idea that large kink points attract significant bunching, and vice versa. The effects, however, are more pronounced among the self-employed than among wage-earners, since the former can more readily change their incomes. Saez (2010) finds evidence of bunching at the first kink point of the US income tax schedule, which is where tax liability starts and is a highly visible point. However, he finds no evidence of bunching at the other kink points, even when they are large. Chetty, et al. (2011a) examine the Danish income tax schedule, and find substantial bunching at the large kink points and little bunching at the smaller kink points. The study by le Maire and Schjerning (2013) also finds bunching around the largest kink points of the Danish income tax schedule. Bastani and Selin (2014) find significant bunching at large kink points of the Swedish income tax schedule among the self-employed, but not among wage-earners.

4 Numerical Examples

To further illustrate the characteristics of locally-optimal income tax schedules, we present two numerical examples. First, we calibrate the model using, where possible, empirically-plausible parameter values. The utility function (2.1) is assumed to take the specific form:

$$x_i - \frac{l_i^{1+\gamma}}{1+\gamma} \tag{4.1}$$

with $\gamma = 2$ as this generates a labour supply elasticity of 0.5, which is in line with empirical estimates (e.g., Chetty et al. 2011b).

The OECD (2018) reports that the ratio of earnings of the fifth income decile to the first income decile averages 1.8, while that for the ninth income decile to the first income decile averages 3.9. We therefore normalise type 1's wage to unity ($w_1 = 1$), and correspondingly set type 3's wage at $w_3 = 1.8$ and type 5's wage at $w_5 = 3.9$. (The wages of the type 2 and type 4 individuals will be chosen such that these individuals are just constrained to earn $\widehat{y}$ and $\overline{y}$, respectively.) There is no empirical guidance as to the welfare weights. Therefore, the weights are arbitrarily chosen to decline linearly and sum to one, i.e., $\pi_1 = 0.30$, $\pi_2 = 0.25$, $\pi_3 = 0.20$, $\pi_4 = 0.15$, and $\pi_5 = 0.10$. The government's exogenous revenue requirement, g, is chosen such that $g / \sum_{i=1}^{5} n_i y_i = 0.4$, which reflects the observation that government spending as a share of GDP averages around 40% in OECD countries. Finally, OECD (2017) data show that approximately 22% of working-age individuals have less than upper-secondary education, 44% have obtained upper-secondary education, and 34% have obtained tertiary education. Consistent with those numbers, we consider two population distributions. First, we present a locally-optimal income tax schedule with few individuals at the kink points, which is therefore close to being linear. In this example, we assume that 21.5% of individuals are low-skill ($n_1 = 0.215$), 45% are middle-skill ($n_3 = 0.45$), and 31.5% are high-skill ($n_5 = 0.315$). The remaining 2% are evenly distributed between type 2 and type 4 individuals ($n_2 = 0.01$ and $n_4 = 0.01$). Second, we present a locally-optimal income tax schedule with many individuals at the kink points, which therefore departs substantially from linearity. In this case, 20% of individuals are low-skill ($n_1 = 0.2$), 40% are middle-skill ($n_3 = 0.4$), and 30% are high-skill ($n_5 = 0.3$). The remaining 10% are split between type 2 and type 4 individuals ($n_2 = 0.05$ and $n_4 = 0.05$).

The locally-optimal income tax schedules under the above parameterisations are shown in Tables 1 and 2. The example when the income tax schedule is approximately linear, shown in Table 1, has marginal tax rates $t_1 = 0.302$, $t_2 = 0.326$, and $t_3 = 0.349$. The example with larger kink points, shown in Table 2, has $t_1 = 0.194$, $t_2 = 0.276$, and $t_3 = 0.374$. Despite the simplicity of the model, the levels of the marginal tax rates are plausible. In both examples, the top marginal tax rate, t_3, is perhaps a little on the low side, but one could argue that the higher top marginal tax rates observed in practice (typically greater than 40%) are designed in part to target the very rich. Under our parameterisations, it can be calculated that $(t_3 - t_2)/t_2 < (t_2 - t_1)/t_1$. This shows that $n_4 \geq n_2$ is not sufficient under a weighted utilitarian social welfare objective (unlike under maxi-min) to establish increasing progressivity of the locally-optimal income tax schedule (cf. equation 3.10).

5 Discussion: Regressive and U-shaped Taxation

Suppose the current income tax schedule is regressive ($t_1 > t_2 > t_3$), as illustrated in Figure 2. The question arises as to whether such an income tax schedule can be locally optimal. It can be seen that the behaviour of the type 1, type 3, and type 5 individuals can be described as under progressive taxation. It can also been seen that no individual would choose to earn income at the kink points. Instead, the 'special' individuals, types 2 and 4, now represent those individuals who are indifferent between earning two levels of income. The population of type 2 individuals will be split between those earning y_2^a

and y_2^b, since they are indifferent between earning those two levels of income. Likewise, the population of type 4 individuals will be split between those earning y_4^b and y_4^c. A similar issue arises if the pattern of marginal tax rates is U-shaped ($t_1 > t_2$, $t_3 > t_2$), as illustrated in Figure 3. The population of type 2 individuals will be split between those earning y_2^a and y_2^b.

Unfortunately, our methodology does not seem suited to analysis of regressive or U-shaped taxation. The problem is that small changes in the tax schedule would cause large jumps in behaviour. For example, consider a small increase in $\widehat{y}$ ($\mathrm{d}\widehat{y} > 0$). The utility of type 2 individuals earning y_2^a would not change, but the utility of type 2 individuals earning y_2^b would fall. Accordingly, all type 2 individuals previously earning y_2^b would jump down to earn y_2^a. These jumps are problematic. However, we do not view this shortcoming of our methodology as a major concern, because the regressive and U-shaped cases do not appear to be empirically relevant. That is, we do not know of any real-world income tax schedules that are well described by regressive or U-shaped taxation.

6 Conclusion

This paper examines the setting of income tax policy from the perspective of governments, taking into account the practical constraints that they face. These are that the government must take the current income tax schedule as the starting point, and any proposed change must be small. Most real-world income tax schedules are progressive and piecewise linear with few tax brackets, though they may differ in the degrees to which they depart from linearity. Our analysis suggests that approximately linear or far from linear income tax schedule can be locally optimal, depending upon the number of individuals attracted to the kink points. Assuming that large kink points attract significant bunching and vice versa, current piecewise linear income tax schedules may be locally optimal despite large differences in their departures from linearity.

7 Appendix

A.1 Individual Behaviour

Under the current income tax system, type 1 individuals have chosen to earn y_1. Thus, they have acted as if maximising (2.1) subject to (2.2). The Lagrangian corresponding to this programme can be written as:

$$\mathcal{L} = x_1 - v\left(\frac{y_1}{w_1}\right) + \lambda\left[a + (1-t_1)y_1 - x_1\right] \tag{A.1}$$

where λ is the Lagrange multiplier. The solution to this programme yields $\lambda = 1$ and $y_1 = y_1(w_1, t_1)$. Let V_1 denote type 1's indirect utility function. By the Envelope Theorem:

$$\frac{\partial V_1}{\partial a} = 1 \text{ and } \frac{\partial V_1}{\partial t_1} = -y_1 < 0 \tag{A.2}$$

Under the current income tax system, type 2 individuals have chosen to earn $\widehat{y}$. Thus, they have acted as if maximising (2.1) subject to (2.2) and the constraint that $y_2 \leq \widehat{y}$. The Lagrangian corresponding to this programme can be written as:

$$\mathcal{L} = x_2 - v\left(\frac{y_2}{w_2}\right) + \lambda\left[a + (1-t_1)y_2 - x_2\right] + \alpha\left[\widehat{y} - y_2\right] \tag{A.3}$$

where λ and α are Lagrange multipliers. Under the assumption that the second constraint is binding ($\alpha > 0$), the solution to this programme yields $\lambda = 1$ and $y_2 = \widehat{y}$. Let V_2 denote type 2's indirect utility function. By the Envelope Theorem:

$$\frac{\partial V_2}{\partial a} = 1,\ \frac{\partial V_2}{\partial t_1} = -\widehat{y} < 0, \text{ and } \frac{\partial V_2}{\partial \widehat{y}} = 1 - t_1 - v'\left(\frac{\widehat{y}}{w_2}\right)\frac{1}{w_2} > 0 \tag{A.4}$$

Under the current income tax system, type 3 individuals have chosen to earn y_3. Thus, they have acted as if maximising (2.1) subject to (2.3). The Lagrangian corresponding to this programme can be written as:

$$\mathcal{L} = x_3 - v\left(\frac{y_3}{w_3}\right) + \lambda\left[a + (t_2 - t_1)\widehat{y} + (1-t_2)y_3 - x_3\right] \tag{A.5}$$

where λ is the Lagrange multiplier. The solution to this programme yields $\lambda = 1$ and $y_3 = y_3(w_3, t_2)$. Let V_3 denote type 3's indirect utility function. By the Envelope Theorem:

$$\frac{\partial V_3}{\partial a} = 1, \ \frac{\partial V_3}{\partial t_1} = -\widehat{y} < 0, \ \frac{\partial V_3}{\partial t_2} = \widehat{y} - y_3 < 0, \text{ and } \frac{\partial V_3}{\partial \widehat{y}} = t_2 - t_1 > 0 \tag{A.6}$$

Under the current income tax system, type 4 individuals have chosen to earn $\overline{y}$. Thus, they have acted as if maximising (2.1) subject to (2.3) and the constraint that $y_4 \leq \overline{y}$. The Lagrangian corresponding to this programme can be written as:

$$\mathcal{L} = x_4 - v\left(\frac{y_4}{w_4}\right) + \lambda\left[a + (t_2 - t_1)\widehat{y} + (1 - t_2)y_4 - x_4\right] + \alpha\left[\overline{y} - y_4\right] \tag{A.7}$$

where λ and α are Lagrange multipliers. Under the assumption that the second constraint is binding ($\alpha > 0$), the solution to this programme yields $\lambda = 1$ and $y_4 = \overline{y}$. Let V_4 denote type 4's indirect utility function. By the Envelope Theorem:

$$\frac{\partial V_4}{\partial a} = 1, \ \frac{\partial V_4}{\partial t_1} = -\widehat{y} < 0, \ \frac{\partial V_4}{\partial t_2} = \widehat{y} - \overline{y} < 0, \ \frac{\partial V_4}{\partial \widehat{y}} = t_2 - t_1 > 0, \text{ and } \frac{\partial V_4}{\partial \overline{y}} = 1 - t_2 - v'\left(\frac{\overline{y}}{w_4}\right)\frac{1}{w_4} > 0 \tag{A.8}$$

Under the current income tax system, type 5 individuals have chosen to earn y_5. Thus, they have acted as if maximising (2.1) subject to (2.4). The Lagrangian corresponding to this programme can be written as:

$$\mathcal{L} = x_5 - v\left(\frac{y_5}{w_5}\right) + \lambda\left[a + (t_2 - t_1)\widehat{y} + (t_3 - t_2)\overline{y} + (1 - t_3)y_5 - x_5\right] \tag{A.9}$$

where λ is the Lagrange multiplier. The solution to this programme yields $\lambda = 1$ and $y_5 = y_5(w_5, t_3)$. Let V_5 denote type 5's indirect utility function. By the Envelope Theorem:

$$\frac{\partial V_5}{\partial a} = 1, \ \frac{\partial V_5}{\partial t_1} = -\widehat{y} < 0, \ \frac{\partial V_5}{\partial t_2} = \widehat{y} - \overline{y} < 0, \ \frac{\partial V_5}{\partial t_3} = \overline{y} - y_5 < 0, \ \frac{\partial V_5}{\partial \widehat{y}} = t_2 - t_1 > 0, \text{ and } \frac{\partial V_5}{\partial \overline{y}} = t_3 - t_2 > 0 \tag{A.10}$$

A.2 Derivation of Equation (3.5)

The gradient of the weighted utilitarian social welfare function is:

$$\nabla W := \langle \sum_{i=1}^{5} \pi_i n_i \frac{\partial V_i}{\partial t_1}, \sum_{i=3}^{5} \pi_i n_i \frac{\partial V_i}{\partial t_2}, \pi_5 n_5 \frac{\partial V_5}{\partial t_3}, \sum_{i=2}^{5} \pi_i n_i \frac{\partial V_i}{\partial \widehat{y}}, \sum_{i=4}^{5} \pi_i n_i \frac{\partial V_i}{\partial \overline{y}}, \sum_{i=1}^{5} \pi_i n_i \rangle \tag{A.11}$$

Using equations (A.2), (A.4), (A.6), (A.8) and (A.10), equation (A.11) can be arranged to yield equation (3.5).

A.3 Motzkin's Theorem of the Alternative

Let A, C, and D be $c_1 \times m$, $c_2 \times m$, and $c_3 \times m$ matrices, respectively, where A is non-vacuous (not all zeros). Then *either*:

$$Az \gg 0^{(c_1)} \qquad Cz \geq 0^{(c_2)} \qquad Dz = 0^{(c_3)} \tag{A.12}$$

has a solution $z \in \mathbb{R}^m$, *or*:

$$b_1 A + b_2 C + b_3 D = 0^{(m)} \tag{A.13}$$

has a solution $b_1 > 0^{(c_1)}$, $b_2 \geq 0^{(c_2)}$, and b_3 sign unrestricted, *but never both.* A proof of Motzkin's Theorem can be found in Mangasarian (1969).

A.4 Proof of Proposition 1

To prove part (i), it is sufficient to show that a solution to the system of equations (3.6) exists. The numerical examples provided in Section 4 show that a solution exists. To prove part (ii), if there exist $\theta \geq 0$ and $\beta > 0$ such that system (3.6) is satisfied, then there must also exist $\theta \geq 0$ and $\beta > 0$ that satisfy (3.6) but with $\beta = 1$. Thus, without loss of generality, we set $\beta = 1$. Expanding (3.6) then yields:

$$\theta \left[n_1 \left(y_1 + t_1 \frac{\partial y_1}{\partial t_1} \right) + (1 - n_1)\widehat{y} \right] - \pi_1 n_1 y_1 - \widehat{y} \sum_{i=2}^{5} \pi_i n_i = 0 \tag{A.14}$$

$$\theta \left[n_3 \left(y_3 + t_2 \frac{\partial y_3}{\partial t_2} \right) - (n_3 + n_4 + n_5)\widehat{y} + (n_4 + n_5)\overline{y} \right] + \pi_3 n_3 (\widehat{y} - y_3) + (\pi_4 n_4 + \pi_5 n_5)(\widehat{y} - \overline{y}) = 0 \tag{A.15}$$

$$\theta \left[n_5 \left(y_5 + t_3 \frac{\partial y_5}{\partial t_3} \right) - n_5 \overline{y} \right] + \pi_5 n_5 (\overline{y} - y_5) = 0 \tag{A.16}$$

$$\theta\left[n_2 t_1 + (n_3+n_4+n_5)(t_1-t_2)\right] + \pi_2 n_2 \left[1 - t_1 - v'(\frac{\widehat{y}}{w_2})\frac{1}{w_2}\right] + (t_2-t_1)\sum_{i=3}^{5}\pi_i n_i = 0 \tag{A.17}$$

$$\theta\left[n_4 t_2 + n_5(t_2-t_3)\right] + \pi_4 n_4 \left[1 - t_2 - v'(\frac{\overline{y}}{w_4})\frac{1}{w_4}\right] + \pi_5 n_5 (t_3-t_2) = 0 \tag{A.18}$$

$$-\theta + \sum_{i=1}^{5}\pi_i n_i = 0 \tag{A.19}$$

Algebraic manipulation of (A.17) and (A.19) yields equation (3.7), and algebraic manipulation of (A.18) and (A.19) yields equation (3.8). ■

A.5 Proof of Proposition 2

The proof of Proposition 2 is analogous to that of Proposition 1, except under a maximin social welfare function we have $\pi_1 = 1$ and $\pi_2 = \pi_3 = \pi_4 = \pi_5 = 0$. Equation (3.10) follows from algebraic manipulation of equation (3.9). ■

References

[1] Apps, P., N. Van Long and R. Rees (2014), "Optimal Piecewise Linear Income Taxation", *Journal of Public Economic Theory*, 16, 523–545.

[2] Bastani, S. and H. Selin (2014), "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule", *Journal of Public Economics*, 109, 36-49.

[3] Blackorby, C. and C. Brett (2000), "Fiscal Federalism Revisited", *Journal of Economic Theory*, 92, 300-317.

[4] Brett, C (1998), "Tax Reform and Collective Family Decision-Making", *Journal of Public Economics*, 70, 425-440.

[5] Chetty, R., J. Friedman, T. Olsen and L. Pistaferri (2011a), "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records", *Quarterly Journal of Economics*, 126, 749-804.

[6] Chetty, R., A. Guren, D. Manoli and A. Weber (2011b), "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins", *American Economic Review*, 101, 471-475.

[7] Dahlby, B (1998), "Progressive Taxation and the Social Marginal Cost of Public Funds", *Journal of Public Economics*, 67, 105-122.

[8] Diamond, P (1998), "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates", *American Economic Review*, 88, 83-95.

[9] Diewert, W (1978), "Optimal Tax Perturbations", *Journal of Public Economics*, 10, 139-177.

[10] Duclos, J., P. Makdissi and Q. Wodon (2008), "Socially Improving Tax Reforms", *International Economic Review*, 49, 1505-1537.

[11] Feldstein, M (1976), "On the Theory of Tax Reform", *Journal of Public Economics*, 6, 77-104.

[12] Fleurbaey, M (2006), "Is Commodity Taxation Unfair?", *Journal of Public Economics*, 90, 1765-1787.

[13] Guesnerie, R (1977), "On the Direction of Tax Reform", *Journal of Public Economics*, 7, 179-202.

[14] Guesnerie, R (1995), *A Contribution to the Pure Theory of Taxation*, Cambridge University Press.

[15] Hatta, T (1977), "A Theory of Piecemeal Policy Recommendations", *Review of Economic Studies*, 44, 1-21.

[16] Konishi, H (1995), "A Pareto-Improving Commodity Tax Reform under a Smooth Nonlinear Income Tax", *Journal of Public Economics*, 56, 413-446.

[17] Krause, A (2007), "Generational Incidence of Savings Taxation versus Capital-Income Taxation", *International Journal of Economic Theory*, 3, 113-129.

[18] Krause, A (2009), "A General Equilibrium Analysis of the Laffer Argument", *Social Choice and Welfare*, 33, 601-615.

[19] le Maire, D. and B. Schjerning (2013), "Tax Bunching, Income Shifting and Self-Employment", *Journal of Public Economics*, 107, 1-18.

[20] Mangasarian, O (1969), *Nonlinear Programming*, McGraw-Hill.

[21] Mankiw, G., M. Weinzierl and D. Yagan (2009), "Optimal Taxation in Theory and Practice", *Journal of Economic Perspectives*, 23, 147-174.

[22] Mirrlees, J (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, 38, 175-208.

[23] Murty, S. and R. Russell (2005), "Externality Policy Reform: A General Equilibrium Analysis", *Journal of Public Economic Theory*, 7, 117-150.

[24] Myles, G (1995), *Public Economics*, Cambridge University Press.

[25] OECD (2017), *Education at a Glance*, OECD publications.

[26] OECD (2018), *Employment and Labour Market Statistics*, OECD publications.

[27] Saez, E (2001), "Using Elasticities to Derive Optimal Income Tax Rates", *Review of Economic Studies*, 68, 205-229.

[28] Saez, E (2010), "Do Taxpayers Bunch at Kink Points?", *American Economic Journal: Economic Policy*, 2, 180-212.

[29] Sheshinski, E (1989), "Note on the Shape of the Optimum Income Tax Schedule", *Journal of Public Economics*, 40, 201-215.

[30] Slemrod, J., S. Yitzhaki, J. Mayshar and M. Lundholm (1994), "The Optimal Two-Bracket Linear Income Tax", *Journal of Public Economics*, 53, 269–290.

[31] Weymark, J (1979), "A Reconciliation of Recent Results in Optimal Taxation", *Journal of Public Economics*, 7, 171-190.

FIGURE 1

Progressive Piecewise Linear Income Taxation

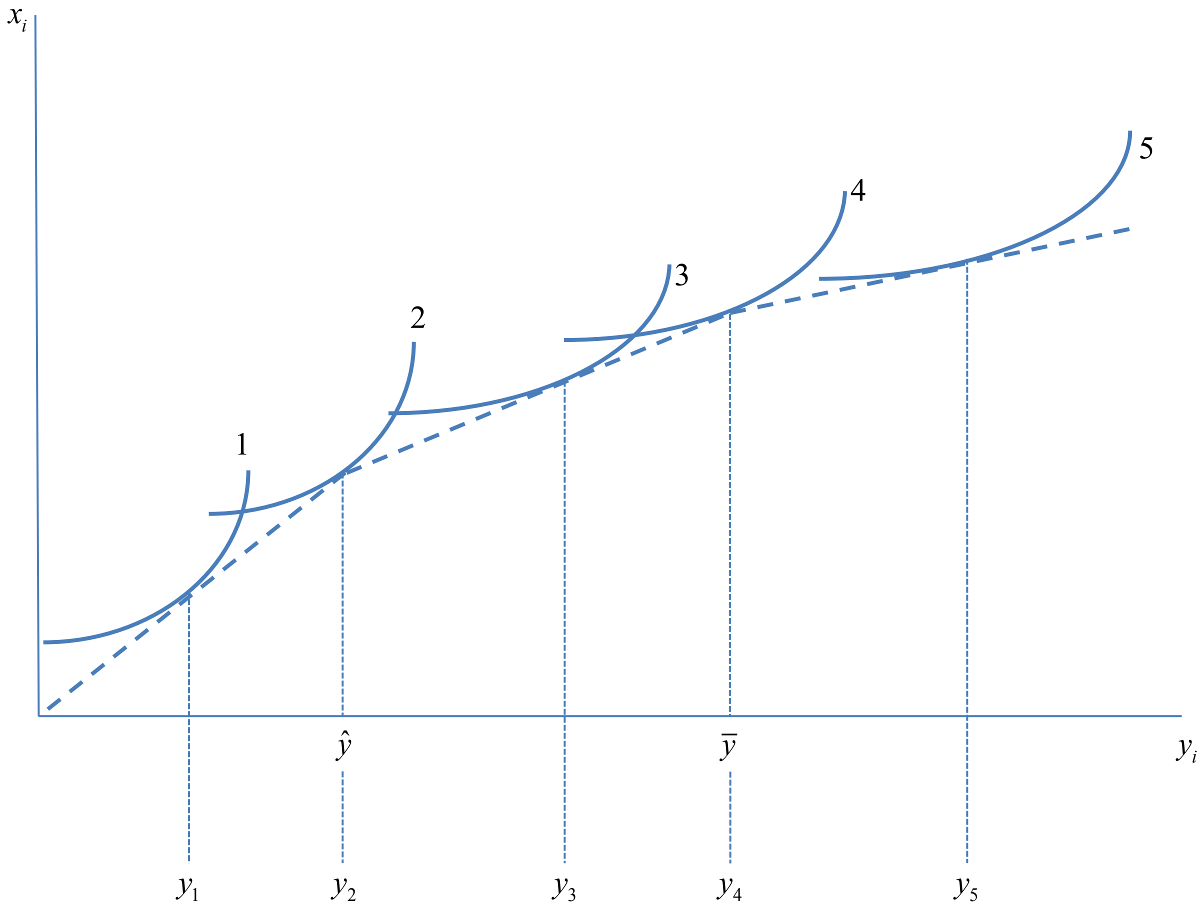

TABLE 1

Numerical Example 1

A Locally Optimal Progressive Three-Bracket Piecewise Linear Income Tax Schedule

		Model Parameters			
n_1	0.215	π_1	0.300	w_1	1.000
n_2	0.010	π_2	0.250	w_3	1.800
n_3	0.450	π_3	0.200	w_5	3.900
n_4	0.010	π_4	0.150	γ	2.000
n_5	0.315	π_5	0.100		
		Tax Parameters			
a	–0.228	$\hat{y}$	1.147	$\bar{y}$	2.706
t_1	0.302	t_2	0.326	t_3	0.349
		Other Variables			
y_1	0.836	$y_2 = \hat{y}$	1.147	w_2	1.242
y_3	1.982	$y_4 = \bar{y}$	2.706	w_4	2.221
y_5	6.216	$g/\sum n_i y_i$	0.400	g	1.227

TABLE 2

Numerical Example 2

A Locally Optimal Progressive Three-Bracket Piecewise Linear Income Tax Schedule

Model Parameters					
n_1	0.200	π_1	0.300	w_1	1.000
n_2	0.050	π_2	0.250	w_3	1.800
n_3	0.400	π_3	0.200	w_5	3.900
n_4	0.050	π_4	0.150	γ	2.000
n_5	0.300	π_5	0.100		
Tax Parameters					
a	–0.344	$\hat{y}$	1.085	$\overline{y}$	2.243
t_1	0.194	t_2	0.276	t_3	0.374
Other Variables					
y_1	0.898	$y_2 = \hat{y}$	1.085	w_2	1.142
y_3	2.055	$y_4 = \overline{y}$	2.243	w_4	1.914
y_5	6.092	$g/\sum n_i y_i$	0.400	g	1.198

FIGURE 2

Regressive Piecewise Linear Income Taxation

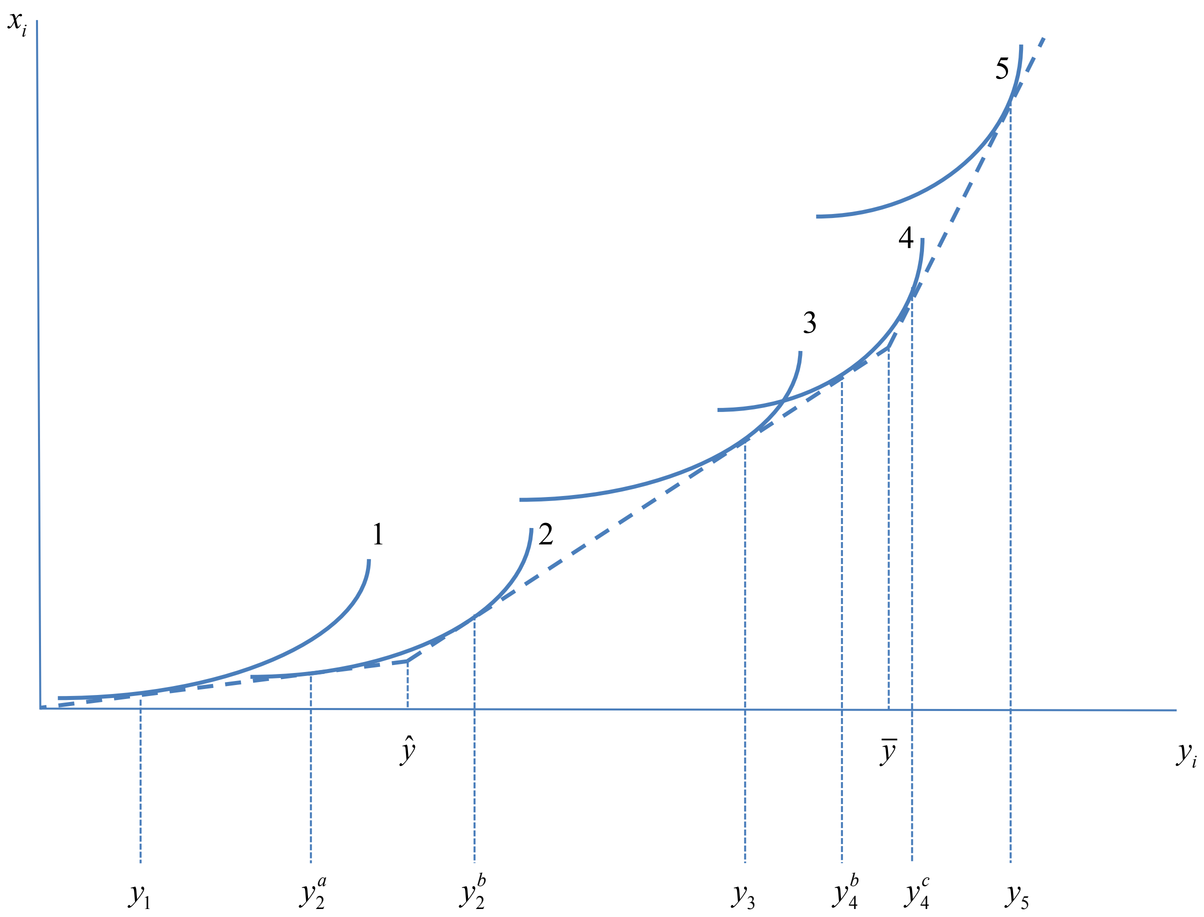

FIGURE 3

U-shaped Piecewise Linear Income Taxation