# Dynamic Income Taxation without Commitment: Comparing Alternative Tax Systems

Jang-Ting Guo*
University of California, Riverside

Alan Krause†
University of York

29 June 2010

**Abstract**

This paper addresses the question as to whether it is optimal to use separating or pooling nonlinear income taxation, or to use linear income taxation, when the government cannot commit to its future tax policy. We consider both two-period and infinite-horizon settings. Under empirically plausible parameter values, separating income taxation is optimal in the two-period model, whereas linear income taxation is optimal when the time horizon is infinite. The welfare effects of varying the discount rate, the degree of wage inequality, and the population of high-skill workers are also explored. For realistic changes in these parameters, separating income taxation remains optimal in the two-period formulation, and linear income taxation remains optimal in the infinite-horizon model.

*Keywords*: Dynamic Income Taxation; Commitment.

*JEL* Classifications: H21, H24.

---

*Department of Economics, 4123 Sproul Hall, University of California, Riverside, CA, 92521, U.S.A., Phone: 1-951-827-1588, Fax: 1-951-827-5685, E-mail: guojt@ucr.edu.

†Corresponding Author. Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K., Phone: 44-(0)1904-433-572, Fax: 44-(0)1904-433-759, E-mail: ak519@york.ac.uk.

# 1 Introduction

Traditionally, macro-style analyses of taxation have used dynamic models, but the common assumption that all individuals are identical rules out a redistributive role for tax policy. On the other hand, micro-style analyses of taxation typically use models with heterogeneous agents, which allows for redistributive concerns, but these models tend to be static which rules out intertemporal considerations. In recent years, a literature known as the "new dynamic public finance" has emerged that seeks to unite the macro- and micro-style approaches by extending the Mirrlees [1971] model of optimal nonlinear income taxation to a dynamic setting.[1] For the most part, this literature has maintained the Mirrlees assumption that there is a continuum of skill types, and it has assumed an infinite time horizon and that future wages are determined by random productivity shocks. Accordingly, the complexity of these models has led most to make the simplifying assumption that the government can commit to its future tax policy. Specifically, the government cannot use skill-type information revealed in earlier periods to redesign the tax system and achieve a better allocation in latter periods.[2]

The commitment assumption might be criticised as being inconsistent with the micro-foundations of the Mirrlees model. In the Mirrlees model, the government cannot observe each individual's skill type, which is the reason it must use (the second-best) incentive-compatible taxation. But such taxation in earlier periods of a dynamic Mirrlees model results in skill-type information being revealed to the government, which would then enable it to implement (the first-best) personalised lump-sum taxes in latter periods. Thus ruling out lump-sum taxation in a dynamic Mirrlees model via a commitment assumption might be considered ad hoc, in much the same way as ruling out lump-

[1]Examples of this literature include Kocherlakota [2005], Albanesi and Sleet [2006], and Werning [2007], among others. Surveys of the new dynamic public finance literature are provided by Golosov, et al. [2006] and Golosov, et al. [2010]. For a textbook treatment of the new dynamic public finance, see Kocherlakota [2010].

[2]Important exceptions that relax the commitment assumption include Farhi and Werning [2008] and Acemoglu, et al. [2008, 2010]. The latter two papers, in particular, are concerned with the revelation and use of skill-type information, but where politicians may use this information partly for their own benefit, rather than only to maximise social welfare. Their analyses are therefore mostly positive in nature, while ours is purely normative.

sum taxation in representative-agent models is considered somewhat artificial.[3] The commitment assumption has also been criticised as being unrealistic, since the present government cannot easily impose constraints on the tax policies of future governments.[4]

The well-known problem with relaxing the commitment assumption is that the revelation principle may no longer hold. That is, it may no longer be social-welfare maximising for the government to design a (separating) nonlinear income tax system in which individuals are willing to reveal their skill types. Instead, it may be optimal to pool the individuals so that skill-type information is not revealed. Similarly, it may be optimal to simply use a linear income tax scheme. Little is known as to under what conditions separating, pooling, or linear income taxation is most desirable from the perspective of maximising social welfare. Roberts [1984] concludes that if the time horizon is infinite and there is no discounting, separation never occurs. The intuition is fairly straightforward: if high-skill individuals live forever, they will forever face personalised lump-sum taxation if they reveal their type. Moreover, since they do not discount the future, they cannot be compensated in the present for the ever-lasting personalised lump-sum taxation they would face after revealing their type. Hence separation is not possible. Berliant and Ledyard [2005] examine a two-period model with discounting. They conclude that separation occurs provided the discount rate is high. The intuition is again fairly straightforward: if high-skill individuals are not too concerned about their future welfare, there exists a relatively low level of compensation that they can be given in period 1 for revealing their type and facing personalised lump-sum taxation in period 2. In this case, separation is not too costly from a social-welfare point of view, and is therefore desirable.

The assumption made by Roberts [1984] that there is no discounting is extreme, and in Berliant and Ledyard [2005] it is not clear whether the "high" discount rate that their conclusion requires is empirically plausible. Also, Roberts [1984] and Berliant and

[3] Indeed, one of the motivations behind the new dynamic public finance literature is to remove the need for ad hoc constraints on the tax instruments available to the government, which must be imposed in standard macro-style dynamic models. See Golosov, et al. [2006] for further discussion.

[4] To be fair, one could argue in favour of the commitment assumption on the basis that real-world tax systems are not frequently redesigned. Gaube [2007], for example, makes this argument.

Ledyard [2005] do not consider the effects of other parameters on the relative desirability of separating taxation.[5] The main objective of this paper is to further investigate under what conditions separating, pooling, or linear income taxation is most desirable. To this end, we use the often-employed two-type version of the Mirrlees model introduced by Stiglitz [1982], but extend it to two-period and infinite-horizon settings. We further assume that preferences take the analytically tractable quasi-linear form, which allows us to obtain a complete characterisation of the solution to the optimal taxation problem and conduct numerical simulations in each of the separating, pooling, and linear income tax cases. Our main results can be summarised as follows. For empirically plausible values of the model's parameters, separating income taxation is optimal in the two-period model, whereas linear income taxation is optimal in the infinite-horizon model. We then examine how the relative desirability of each tax system is affected by changes in some key parameters, namely, the discount rate, the degree of wage inequality, and the population of high-skill workers. For reasonable changes in these parameters, it is shown that separating income taxation remains optimal in the two-period model, and linear income taxation remains optimal in the infinite-horizon model. Separating income taxation increases its advantage in the two-period model when the discount rate, the degree of wage inequality, and/or the population of high-skill workers rises. Linear income taxation increases its advantage in the infinite-horizon model when the degree of wage inequality and/or the population of high-skill workers rises. Finally, separating income taxation is not feasible in the infinite-horizon model for all realistic values of the parameters.

The remainder of the paper is organised as follows. Section 2 describes the analytical framework that we consider. Section 3 describes the structure of separating, pooling, and linear income taxation in the two-period model, and then discusses the results of our numerical simulations. Section 4 describes the structure of each tax system when the

[5] Similar nonlinear income tax models without commitment have been used by Apps and Rees [2006], Bisin and Rampini [2006], Brett and Weymark [2008], Krause [2009], and Guo and Krause [2010], among others. These papers all assume a two-period time horizon and that there are only two skill types. However, none of these papers address the issue of whether separating or pooling is optimal, with most simply considering in turn separating and pooling taxation.

model is extended to an infinite horizon, and then discusses the corresponding numerical simulation results. Section 5 concludes, with two appendices containing some additional mathematical details.

## 2 Preliminaries

We first consider an economy that lasts for two periods, and then consider an extension to an infinite-horizon setting. There is a unit measure of individuals who live for the duration of the economy, with a proportion $\phi \in (0, 1)$ being high-skill workers and the remaining $(1 - \phi)$ being low-skill workers. The wage rates of the high-skill and low-skill types are denoted by $w_H$ and $w_L$, respectively, where $w_H > w_L$. Wages are assumed to remain constant through time. Both types have the same preferences over consumption and labour in each period, which are represented by the analytically tractable quasi-linear utility function:

$$\alpha c_i^t - \frac{1}{1+\gamma}(l_i^t)^{1+\gamma} \tag{2.1}$$

where $c_i^t$ denotes type $i$'s consumption in period $t$, $l_i^t$ denotes type $i$'s labour supply in period $t$, while $\alpha > 0$ and $\gamma > 0$ are preference parameters. All individuals discount the future using the discount factor $\delta = \frac{1}{1+r}$, where $r > 0$ is the discount rate. Type $i$'s pre-tax income in period $t$ is denoted by:

$$y_i^t = w_i l_i^t \tag{2.2}$$

For simplicity we assume there are no savings, which implies that $y_i^t - c_i^t$ is equal to total taxes paid (or transfers received) by a type $i$ individual in period $t$.

The government seeks to maximise social welfare over the duration of the economy, which is assumed measurable by a utilitarian social welfare function weighted towards low-skill individuals. The government will therefore be using its taxation powers to redistribute from high-skill individuals to low-skill individuals. However, the government cannot implement (the first-best) personalised lump-sum taxes in each period, since following the standard practice we assume that each individual's skill type is initially private information. In static models of this kind, it is well known that the best the government can do is to implement (the second-best) incentive-compatible taxation in which each individual is willing to reveal their type (see, e.g., Stiglitz [1982]). But since our model is dynamic and the government cannot commit, each individual knows that if they reveal their type in period 1 they will be subjected to personalised lump-sum taxation thereafter. This implies that high-skill individuals must be offered a relatively favourable tax treatment in period 1 if they are to reveal their type, in order to compensate for the unfavourable tax treatment they will receive in periods $t \geq 2$. From the government's point of view, the lack of redistribution it can undertake in period 1 if skill-type information is to be obtained may be very costly in terms of the level of social welfare attainable. Instead, a higher level of social welfare might be obtained if the government were to pool the individuals in period 1 so that no skill-type information is revealed, even though it is then constrained to use second-best taxation in period 2 in the two-period model or to keep on pooling forever in the infinite-horizon model. Likewise, social welfare could be higher if the government simply used linear income taxation in each period.[6]

As with the individuals, for simplicity we assume that the government cannot save or borrow. Therefore, the only link between periods is the possible revelation and use of skill-type information.

# 3 Two-Period Model

Deciding whether the government should use a nonlinear income tax system that: (1) separates in period 1 and uses first-best taxation in period 2, (2) pools in period 1 and uses second-best taxation in period 2, or (3) simply use linear income taxation in both periods, requires a comparison of social welfare in each case, and such comparisons

[6] If the government uses linear income taxation in period 1, high-skill and low-skill individuals will typically make different consumption/labour choices, thus revealing their skill types. In principle, this would enable the government to implement personalised lump-sum taxes thereafter. Indeed, this is the main issue with which Dillen and Lundholm [1996] are concerned. We do not consider this possibility here, however, since we wish to examine how a simple linear income tax system compares with nonlinear separating and pooling income taxation.

depend upon the values of the model's parameters. Accordingly, in this section we describe in turn the structure of separating, pooling, and linear income tax systems, which then form the basis for social-welfare comparisons made via numerical simulations. Further details on each tax system are provided in Appendix A.

### 3.1 Separation in Period 1 and First-Best Taxation in Period 2

If the tax system in period 1 was designed to separate high-skill and low-skill individuals, the government can use skill-type information revealed in period 1 to implement personalised lump-sum taxes in period 2. In this case, the government's behaviour in period 2 can be described as follows. Choose tax treatments $\langle c_L^2, y_L^2 \rangle$ and $\langle c_H^2, y_H^2 \rangle$ for the low-skill and high-skill individuals, respectively, to maximise:

$$\pi(1-\phi)\left[\alpha c_L^2 - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_L}\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha c_H^2 - \frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma}\right] \tag{3.1}$$

subject to:

$$(1-\phi)(y_L^2 - c_L^2) + \phi(y_H^2 - c_H^2) \geq 0 \tag{3.2}$$

$$c_H^2 = \frac{1-\pi}{\pi} c_L^2 \tag{3.3}$$

where (3.1) is the second-period weighted utilitarian social welfare function, with the utility functions written in terms of the government's choice variables, and $\pi \in (\frac{1}{2}, 1)$ is the weight the government attaches to the welfare of low-skill individuals.[7] Equation (3.2) is the government's second-period budget constraint, while (3.3) is a constraint that is used to determine the first-best consumption levels. As has been noted previously by others, the first-best optimal income tax problem with quasi-linear preferences is generally not completely determinate.[8] However, it is well known that when preferences take the more general additively-separable form, $u(c_i^t) - v(l_i^t)$, with $u(\cdot)$ increasing and strictly concave and $v(\cdot)$ increasing and strictly convex, first-best income taxation under a strict utilitarian (i.e., $\pi = 0.5$) objective gives both types the same level of consumption, but

[7] Since we are working with quasi-linear in consumption preferences, we must assume that $\pi > 0.5$ to ensure that the high-skill type's incentive-compatibility constraint is binding.

[8] See, e.g., Brett and Weymark [2008] and Krause [2009]. Specifically, when the utility function is quasi-linear in consumption (labour), the first-best levels of consumption (pre-tax income) cannot be determined.

the high-skill type is required to work longer.[9] We therefore use this insight to determine the first-best consumption levels with quasi-linear preferences by adding constraint (3.3).[10] The solution to programme (3.1) – (3.3) yields functions for the choice variables that generally depend upon the parameters of the problem, i.e., $c_L^2(\pi, \phi, \alpha, \gamma, w_L, w_H)$, $y_L^2(\cdot)$, $c_H^2(\cdot)$, and $y_H^2(\cdot)$. Substituting these functions into (3.1) yields the level of social welfare in period 2 under first-best taxation, which we denote by $W_F^2(\cdot)$.

All individuals know that if they reveal their skill type in period 1, the government will solve programme (3.1) – (3.3) in period 2. Therefore, in order to induce each individual to reveal their type in period 1, the government chooses tax treatments $\langle c_L^1, y_L^1 \rangle$ and $\langle c_H^1, y_H^1 \rangle$ for the low-skill and high-skill individuals, respectively, to maximise:

$$\pi(1-\phi)\left[\alpha c_L^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_L}\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha c_H^1 - \frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma}\right] \tag{3.4}$$

subject to:

$$(1-\phi)(y_L^1 - c_L^1) + \phi(y_H^1 - c_H^1) \geq 0 \tag{3.5}$$

$$\alpha c_H^1 - \frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} + \delta\left[\alpha c_H^2(\cdot) - \frac{1}{1+\gamma}\left(\frac{y_H^2(\cdot)}{w_H}\right)^{1+\gamma}\right] \geq$$

$$\alpha c_L^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} + \delta\left[\alpha c_L^2(\cdot) - \frac{1}{1+\gamma}\left(\frac{y_L^2(\cdot)}{w_H}\right)^{1+\gamma}\right] \tag{3.6}$$

where (3.4) is first-period social welfare, (3.5) is the government's first-period budget constraint, and (3.6) is the high-skill type's incentive-compatibility constraint. Since skill type is private information in period 1, the government must satisfy incentive-compatibility constraints to ensure that each type chooses their intended tax treatment, rather than mimicking the other type by choosing the other type's tax treatment. However, we omit the low-skill type's incentive-compatibility constraint, as we follow the standard practice of focusing on "redistributive equilibria". That is, we assume that the

[9]This has led some to describe first-best income taxation under a utilitarian objective as Marxist in nature, since it takes from each individual according to their ability and gives to each individual according to their need.

[10]While the addition of constraint (3.3) to make the first-best optimal income tax problem with quasi-linear preferences fully determinate might be considered a little ad hoc, it will be seen that it has no real bearing on our main conclusions.

redistributive goals of the government create an incentive for high-skill individuals to mimic low-skill individuals, but not vice versa. This implies that the high-skill type's incentive-compatibility constraint will bind at an optimum, whereas the low-skill type's incentive-compatibility constraint will be slack.[11] In order to induce high-skill individuals to reveal their type in period 1, the utility they obtain from choosing $\langle c_H^1, y_H^1 \rangle$ in period 1 and thus revealing their type, plus the utility they obtain from $\langle c_H^2(\cdot), y_H^2(\cdot) \rangle$ which they are then forced to accept in period 2, must be greater than or equal to the utility they could obtain by pretending to be low skill by choosing $\langle c_L^1, y_L^1 \rangle$ in period 1, plus the utility they obtain from the low-skill type's tax treatment $\langle c_L^2(\cdot), y_L^2(\cdot) \rangle$ in period 2. That is, if a high-skill individual pretends to be low skill in period 1, they will be treated as such by the government in period 2. The solution to programme (3.4) – (3.6) yields functions for the choice variables $c_L^1(\pi, \phi, \alpha, \gamma, w_L, w_H, \delta)$, $y_L^1(\cdot)$, $c_H^1(\cdot)$, and $y_H^1(\cdot)$. Substituting these functions into (3.4) yields the level of social welfare in period 1 under second-best taxation, which we denote by $W_S^1(\cdot)$. Total social welfare under first-period separation is then equal to $W_S^1(\cdot) + \delta W_F^2(\cdot)$.

### 3.2 Pooling in Period 1 and Second-Best Taxation in Period 2

If the individuals were pooled in the first period, the government cannot distinguish high-skill from low-skill individuals in the second period. It therefore solves a standard Mirrlees/Stiglitz optimal nonlinear income tax problem in period 2. That is, the government chooses tax treatments $\langle c_L^2, y_L^2 \rangle$ and $\langle c_H^2, y_H^2 \rangle$ for the low-skill and high-skill individuals, respectively, to maximise:

$$\pi(1-\phi)\left[\alpha c_L^2 - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_L}\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha c_H^2 - \frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma}\right] \tag{3.7}$$

subject to:

$$(1-\phi)(y_L^2 - c_L^2) + \phi(y_H^2 - c_H^2) \geq 0 \tag{3.8}$$

$$\alpha c_H^2 - \frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma} \geq \alpha c_L^2 - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_H}\right)^{1+\gamma} \tag{3.9}$$

[11] This is what Stiglitz [1982] refers to as the "normal" case and what Guesnerie [1995] refers to as "redistributive equilibria".

where (3.7) is second-period social welfare, (3.8) is the government's second-period budget constraint, and (3.9) is the high-skill type's incentive-compatibility constraint. The solution to programme (3.7)−(3.9) yields functions for the choice variables $c_L^2(\pi, \phi, \alpha, \gamma, w_L, w_H)$, $y_L^2(\cdot)$, $c_H^2(\cdot)$, and $y_H^2(\cdot)$. Substituting these functions into (3.7) yields the level of social welfare in period 2 under second-best taxation, which we denote by $W_S^2(\cdot)$.

If the government decides to pool the individuals in period 1, it chooses a single tax treatment for both types $\langle c^1, y^1 \rangle$ to maximise first-period social welfare:

$$\pi(1-\phi)\left[\alpha c^1 - \frac{1}{1+\gamma}\left(\frac{y^1}{w_L}\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha c^1 - \frac{1}{1+\gamma}\left(\frac{y^1}{w_H}\right)^{1+\gamma}\right] \tag{3.10}$$

subject to the government's first-period budget constraint:

$$y^1 - c^1 \geq 0 \tag{3.11}$$

Since the budget constraint will bind at an optimum, the solution to programme (3.10)−(3.11) will involve $c^1 = y^1 = y^1(\pi, \phi, \alpha, \gamma, w_L, w_H)$. Substituting this function into (3.10) yields the level of social welfare in period 1 under pooling, which we denote by $W_P^1(\cdot)$. Total social welfare under first-period pooling is then equal to $W_P^1(\cdot) + \delta W_S^2(\cdot)$.

### 3.3 Linear Income Taxation in Both Periods

If the government uses linear income taxation in both periods, each individual $i$ will solve the following problem in each period. Choose $c_i^t$ and $l_i^t$ to maximise:

$$\alpha c_i^t - \frac{1}{1+\gamma}\left(l_i^t\right)^{1+\gamma} \tag{3.12}$$

subject to their period $t$ budget constraint:

$$c_i^t \leq a^t + (1-\tau^t) w_i l_i^t \tag{3.13}$$

where the linear income tax system comprises a uniform lump-sum transfer $a^t$ and income tax rate $\tau^t$ in each period. The solution to programme (3.12) − (3.13) will yield the functions $c_i^t(\alpha, \gamma, w_i, a^t, \tau^t)$ and $l_i^t(\cdot)$.

Since the government knows that all individuals will solve programme (3.12) – (3.13) in each period, it designs an optimal linear income tax system in each period by choosing $a^t$ and $\tau^t$ to maximise:

$$\pi(1-\phi)\left[\alpha c_L^t(\cdot) - \frac{1}{1+\gamma}\left(l_L^t(\cdot)\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha c_H^t(\cdot) - \frac{1}{1+\gamma}\left(l_H^t(\cdot)\right)^{1+\gamma}\right] \quad (3.14)$$

subject to:

$$(1-\phi)\tau^t w_L l_L^t(\cdot) + \phi\tau^t w_H l_H^t(\cdot) \geq a^t \quad (3.15)$$

where (3.14) is the weighted utilitarian social welfare function, and (3.15) is the government's period $t$ budget constraint. The solution to programme (3.14) – (3.15) will yield functions $a^t(\pi, \phi, \alpha, \gamma, w_L, w_H)$ and $\tau^t(\cdot)$. These functions can then be used to determine the optimal levels of $c_i^t(\cdot)$ and $l_i^t(\cdot)$, which in turn can be substituted into (3.14) to determine social welfare in each period, which we denote by $W_{LI}^t(\cdot)$. Total social welfare under linear income taxation is then equal to $W_{LI}^1(\cdot) + \delta W_{LI}^2(\cdot)$.

### 3.4 Numerical Simulations

Based on the preceding analytical analyses, it is not possible to rank, from a social-welfare perspective, the relative desirability of the three tax systems. Therefore, this subsection conducts a quantitative welfare comparison. We begin by identifying a set of baseline parameter values that are reasonable. These are presented in Table 1. The parameters are chosen on the following basis. If $\pi = 0.5$, the social welfare function is strictly utilitarian, which is a common assumption in the literature. However, since we assume quasi-linear in consumption preferences, we must assume that $\pi > 0.5$ to ensure that the high-skill type's incentive-compatibility constraint is binding. We therefore set $\pi = 0.55$ so that our weighted utilitarian social welfare function is approximately utilitarian. Walker and Zhu [2008] report that in 2006 university graduates comprised more than 13% of the U.K. workforce, but they also report that the proportion of young people who go on to university is now around 35%. We therefore assume that 25% of individuals are high-skill workers, i.e., we set $\phi = 0.25$. The preference parameters $\alpha$ and $\gamma$ are both set to unity so that the period utility function (2.1) is quadratic in hours worked. We assume an annual discount rate of 4%, which is in line with the real business cycle literature. Since most individuals work for around 40 years of their lives, we take each period to be 20 years in length. An annual discount rate of 4% then corresponds to a 20-year discount factor of $\delta = 0.46$. Fang [2006] and Goldin and Katz [2007] estimate that the college wage premium, i.e., the average difference in the wages of university graduates over high-school graduates, is approximately 60%. We therefore normalise the low-skill type's wage rate to unity ($w_L = 1$), and set the high-skill type's wage rate at $w_H = 1.6$.

For these parameter values, Table 1 shows that separating income taxation is social-welfare maximising, linear income taxation is ranked second, while pooling is third. Social welfare in period 1 is actually lower under separating income taxation than under linear income taxation, because in period 1 under separation high-skill individuals have to be compensated for the first-best tax treatment they will face in period 2. Indeed, high-skill individuals face a negative average tax rate in period 1. However, social welfare in period 2 under separation is significantly higher, since first-best taxation is far superior to a simple linear income tax. Pooling yields the lowest level of social welfare, even though it allows second-best taxation to be used in period 2, which is better than linear income taxation. However, pooling in period 1 is very costly, as reflected in the low level of social welfare.

Figure 1 shows the effects of reasonable variations in the relative size of the high-skill population ($\phi$), the discount rate ($r$), and the wage premium ($w_H/w_L$), whilst holding all other parameters at their baseline levels. The social-welfare ranking of separating income taxation, linear income taxation, and pooling remains unchanged for the variations considered. Social welfare is increasing under each tax system as $\phi$ increases, which simply reflects the fact that society is better-off with a larger population of high-skill workers. Separation increases its advantages over linear income taxation and pooling as $\phi$ increases. An increase in $\phi$ implies that high-skill individuals receive a greater weight in the social welfare function, which means redistribution in period 2 under first-best taxation becomes less severe. This in turn implies that high-skill individuals require less compensation in period 1 to reveal their type, thus making separation more attractive. Linear income taxation also increases its advantage over pooling as $\phi$ increases.

Increases in $\phi$ exacerbate the redistributive inefficiency of pooling in period 1, since the greater weight high-skill individuals receive in the social welfare function, combined with the pooling restriction that both types receive the same allocation, imply that high-skill individuals are made better-off and low-skill individuals are made worse-off in period 1. This inefficiency is partly reversed in period 2 when nonlinear income taxation is used after pooling, since nonlinear income taxation is more effective in achieving redistribution than linear income taxation. But the advantage is not sufficient to overcome the increased inefficiency of pooling in the first period.

As the discount rate $r$ increases, social welfare declines under each tax system simply because less weight is attached to the value of second-period social welfare. Higher values of $r$ increase the advantages that separation has over linear income taxation and pooling, albeit only slightly. As $r$ increases, high-skill individuals become less concerned with the low level of utility they obtain under first-best taxation in period 2. Accordingly, the utility they require in period 1 as compensation for revealing their type decreases, making separation less costly. Increases in $r$ also make linear income taxation more attractive than pooling. Since pooling in period 1 is less efficient than linear income taxation, but nonlinear income taxation in period 2 is more efficient than linear income taxation, increases in $r$ make pooling in period 1 along with nonlinear income taxation in period 2 less attractive because an increase in $r$ implies a relatively higher concern for first-period social welfare and a lower concern for second-period social welfare.

As the wage premium $w_H/w_L$ increases, social welfare increases under each tax system. Since $w_L$ is normalised to unity, an increase in $w_H/w_L$ effectively means an increase in the economy's endowments, hence social welfare also rises. Separation increases its advantages over linear income taxation and pooling as $w_H/w_L$ increases. Given the government's redistributive concerns, linear income taxation in both periods or pooling in period 1 along with second-best taxation in period 2 are not as powerful as separating the individuals in period 1 and then being able to use first-best taxation in period 2. Moreover, the relative desirability of separating taxation is naturally increasing in the degree of wage inequality, since the need for redistribution increases. As $w_H/w_L$ increases, linear income taxation is also increasingly preferred to pooling. On the one hand, an increase in wage inequality exacerbates the inefficiency of pooling relative to linear income taxation in period 1, but on the other hand higher wage inequality increases the advantage nonlinear income taxation has over linear income taxation in period 2. However, on balance our numerical simulations indicate that linear income taxation is increasingly preferred over pooling as $w_H/w_L$ increases.

# 4 Infinite-Horizon Model

In this section, we describe how the general structure of separating, pooling, and linear income taxation changes when the model is extended from two periods to an infinite-horizon setting. Further details are provided in Appendix B.

### 4.1 Separation in Period 1 and First-Best Taxation Thereafter

If the individuals were separated in period 1, the government can implement personalised lump-sum taxation from period 2 onwards. That is, the government will solve programme (3.1) − (3.3) in periods $2, ..., \infty$. Let $u^t_{iF}(\pi, \phi, \alpha, \gamma, w_L, w_H)$ denote the utility type $i$ obtains under first-best taxation in each period, and let $W^t_F(\cdot)$ denote the level of social welfare under first-best taxation in each period.

If skill-type information is revealed in period 1, everyone knows that the government will solve programme (3.1) − (3.3) in periods $2, ..., \infty$. Therefore, in order to induce individuals to reveal their types in period 1, the government chooses tax treatments $\langle c^1_L, y^1_L \rangle$ and $\langle c^1_H, y^1_H \rangle$ for the low-skill and high-skill types, respectively, to maximise:

$$\pi(1-\phi)\left[\alpha c^1_L - \frac{1}{1+\gamma}\left(\frac{y^1_L}{w_L}\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha c^1_H - \frac{1}{1+\gamma}\left(\frac{y^1_H}{w_H}\right)^{1+\gamma}\right] \tag{4.1}$$

subject to:

$$(1-\phi)(y^1_L - c^1_L) + \phi(y^1_H - c^1_H) \geq 0 \tag{4.2}$$

$$\alpha c^1_H - \frac{1}{1+\gamma}\left(\frac{y^1_H}{w_H}\right)^{1+\gamma} + \sum_{t=2}^{\infty}\delta^{t-1}u^t_{HF}(\cdot) \geq \alpha c^1_L - \frac{1}{1+\gamma}\left(\frac{y^1_L}{w_H}\right)^{1+\gamma} + \sum_{t=2}^{\infty}\delta^{t-1}u^t_{LF}(\cdot) \tag{4.3}$$

where (4.1) is first-period social welfare, (4.2) is the government's first-period budget constraint, and (4.3) is the high-skill type's incentive-compatibility constraint. If high-skill individuals are willing to reveal their type, the utility they obtain from choosing $\langle c_H^1, y_H^1 \rangle$ in period 1 and thus revealing their type, plus the discounted sum of utilities they obtain under first-best taxation from period 2 onwards, must be greater than or equal to the utility they could obtain by pretending to be low skill. The solution to programme (4.1) – (4.3) yields functions for the choice variables $c_L^1(\pi, \phi, \alpha, \gamma, w_L, w_H, \delta)$, $y_L^1(\cdot)$, $c_H^1(\cdot)$, and $y_H^1(\cdot)$. Substituting these functions into (4.1) yields the level of social welfare in period 1 under second-best taxation, which we denote by $W_S^1(\cdot)$. Total social welfare under separation is then equal to $W_S^1(\cdot) + \sum_{t=2}^{\infty} \delta^{t-1} W_F^t(\cdot)$.

### 4.2 Pooling in Each Period

In the two-period model, the government can solve a standard Mirrlees/Stiglitz optimal nonlinear income tax problem in period 2 after pooling in period 1, because there are no later periods in which the government can take advantage of skill-type information revealed in period 2. In the infinite-horizon model, however, there is no last period in which the government can solve a standard nonlinear income tax problem. Therefore, pooling in the infinite-horizon model means pooling in every period, i.e., the government solves programme (3.10) – (3.11) in each period. Total social welfare under pooling is therefore equal to $\sum_{t=1}^{\infty} \delta^{t-1} W_P^t(\pi, \phi, \alpha, \gamma, w_L, w_H)$, where $W_P^t(\cdot)$ is the level of social welfare associated with programme (3.10) – (3.11).

### 4.3 Linear Income Taxation in Each Period

If the government uses linear income taxation in each period, individuals will solve programme (3.12) – (3.13) in each period and the government will solve programme (3.14) – (3.15) in each period. Total social welfare under linear income taxation will therefore be equal to $\sum_{t=1}^{\infty} \delta^{t-1} W_{LI}^t(\pi, \phi, \alpha, \gamma, w_L, w_H)$, where $W_{LI}^t(\cdot)$ is the level of social welfare associated with programme (3.14) – (3.15).

### 4.4 Numerical Simulations

Table 2 presents baseline parameter values for the infinite-horizon model. These are identical to those for the two-period model, except following convention we now take each period to be one year in length. This implies that an annual discount rate of 4% corresponds to a one-year discount factor of $\delta = 0.96$. For the baseline parameter values, separating income taxation is not feasible. That is, the compensation high-skill individuals would require for revealing their type in period 1 and forever-after facing personalised lump-sum taxation is so large that it would necessitate that low-skill individuals face an average tax rate in period 1 of more than 100%. The intuition for this result is similar to that for the result of Roberts [1984] that separation never occurs if there is no discounting and the time horizon is infinite. In order for separation to be feasible in our infinite-horizon model, the annual discount rate would have to be at least 37%, which seems to be unreasonable. Therefore, the only options available to the government in the infinite-horizon model are to pool the individuals in every period or use linear income taxation in every period. But pooling is extreme in that it imposes the same consumption/pre-tax income allocation on both types, whereas the allocations differ under linear income taxation even though both types face the same income tax rate and lump-sum transfer. Therefore, linear income taxation is social-welfare maximising in the infinite-horizon model.

Figure 2 shows the effects of varying the high-skill population $\phi$, the discount rate $r$, and the wage premium $w_H/w_L$ on the relative desirability of linear income taxation and pooling. For increases in $\phi$ and $w_H/w_L$, linear income taxation increases its advantage over pooling. The intuition is similar to that discussed for the first period of our two-period model, and therefore is not repeated here. As $r$ increases, the social-welfare gap between linear income taxation and pooling appears to narrow, but proportionally the advantage linear income taxation has over pooling remains constant. This is because the same allocation is implemented under linear income taxation in each period, as well as under pooling in each period. Therefore, a higher value of $r$ simply means that a lower discount factor is used to sum the infinite social-welfare streams under each tax system.

# 5 Concluding Comments

This paper has addressed the question as to whether it is optimal to use separating or pooling nonlinear income taxation, or to use linear income taxation, when the government cannot commit to its future tax policy. The question is an important one in light of the new dynamic public finance literature, once the commitment assumption is relaxed.

We have shown that separating income taxation is optimal in the two-period model, whereas linear income taxation is optimal when the time horizon is infinite. These results, however, are dependent upon the use of empirically plausible values of the model's parameters, since it is straightforward to show that there exist other (albeit unrealistic) sets of parameter values under which each tax system is optimal. We have also examined how the relative desirability of each tax system is affected by reasonable changes in some key parameters around their baseline values.

Deciding whether separating, pooling, or linear income taxation is most desirable requires a comparison of social welfare in each case. In order to make such comparisons, we have used an analytically tractable preference formulation that yields a complete solution to the optimal tax problem under each tax system. The question remains as to how dependent our results are on the specifics of the model.[12] Based on the nature of the intuition driving our results, we conjecture that many would hold-up in more general settings. Nevertheless, one way to test the robustness of our results, whilst maintaining tractability, would be to change the form of the utility function from quasi-linear in consumption to quasi-linear in labour. However, any form of quasi-linearity will render the first-best optimal tax problem that must be solved in the separating case indeterminate unless something akin to constraint (3.3) is included, and it does not seem clear as to what "reasonable" constraint one could add to determine the first-best levels of pre-tax income. Better still, one could work with the more general additively-separable form of the utility function, but with these preferences it is likely to be extremely difficult to obtain a detailed solution to the optimal tax problem under each tax system.

The only dynamic link in our model is the (possible) revelation and use of skill-type information. It would be interesting to extend the model to allow for savings and borrowings by individuals and, in particular, the government. Allowing the government to transfer resources over time could help it overcome some of the inefficiencies associated with each tax system. However, such efforts by the government might be undermined

[12] That said, we note that constraint (3.3) which is introduced for the sole purpose of making the first-best optimal tax problem with quasi-linear in consumption preferences fully determinate is innocuous, in that it has no real influence on our results.

by individual savings behaviour. These seem interesting issues for future research.

# 6 Appendix A: Two-Period Model

**Separating Income Taxation: Further Details**

To solve programme (3.1) – (3.3), first substitute equation (3.3) into (3.1) and (3.2). The first-order conditions on $c_L^2$, $y_L^2$, and $y_H^2$ can then be written as, respectively:

$$\pi(1-\phi)\alpha + \frac{(1-\pi)\phi\alpha(1-\pi)}{\pi} - \lambda^2(1-\phi) - \frac{\lambda^2\phi(1-\pi)}{\pi} = 0 \tag{A.1}$$

$$-\pi(1-\phi)\left(\frac{y_L^2}{w_L}\right)^{\gamma}\frac{1}{w_L} + \lambda^2(1-\phi) = 0 \tag{A.2}$$

$$-(1-\pi)\phi\left(\frac{y_H^2}{w_H}\right)^{\gamma}\frac{1}{w_H} + \lambda^2\phi = 0 \tag{A.3}$$

where $\lambda^2$ is the multiplier on the government's second-period budget constraint (3.2). Equation (A.1) can be manipulated to yield:

$$\lambda^2 = \frac{\alpha\left[\pi(1-\phi)\pi + (1-\pi)\phi(1-\pi)\right]}{\pi(1-\phi)+(1-\pi)\phi} \tag{A.4}$$

while equations (A.2) and (A.3) can be manipulated to yield, respectively:

$$y_L^2 = \left[\frac{\lambda^2(w_L)^{1+\gamma}}{\pi}\right]^{\frac{1}{\gamma}} \tag{A.5}$$

$$y_H^2 = \left[\frac{\lambda^2(w_H)^{1+\gamma}}{1-\pi}\right]^{\frac{1}{\gamma}} \tag{A.6}$$

After substituting (3.3) into (3.2), we obtain:

$$c_L^2 = \frac{\pi\left[(1-\phi)y_L^2 + \phi y_H^2\right]}{\pi(1-\phi)+(1-\pi)\phi} \tag{A.7}$$

Finally, (3.3) can then be used to determine $c_H^2$.

The relevant first-order conditions corresponding to programme (3.4) – (3.6) can be written as:

$$\pi(1-\phi)\alpha - \lambda^1(1-\phi) - \theta_H\alpha = 0 \tag{A.8}$$

$$-\pi(1-\phi)\left(\frac{y_L^1}{w_L}\right)^\gamma \frac{1}{w_L} + \lambda^1(1-\phi) + \theta_H\left(\frac{y_L^1}{w_H}\right)^\gamma \frac{1}{w_H} = 0 \tag{A.9}$$

$$(1-\pi)\phi\alpha - \lambda^1\phi + \theta_H\alpha = 0 \tag{A.10}$$

$$-(1-\pi)\phi\left(\frac{y_H^1}{w_H}\right)^\gamma \frac{1}{w_H} + \lambda^1\phi - \theta_H\left(\frac{y_H^1}{w_H}\right)^\gamma \frac{1}{w_H} = 0 \tag{A.11}$$

where $\lambda^1$ is the multiplier on the government's first-period budget constraint (3.5), and $\theta_H$ is the multiplier on the high-skill type's incentive-compatibility constraint (3.6). Equations (A.8) and (A.10) can be combined to yield:

$$\lambda^1 = \alpha\left[\pi(1-\phi) + (1-\pi)\phi\right] \tag{A.12}$$

$$\theta_H = \frac{\lambda^1\phi}{\alpha} - (1-\pi)\phi \tag{A.13}$$

Equations (A.9) and (A.11) can be manipulated to yield, respectively:

$$y_L^1 = \left[\frac{\lambda^1(1-\phi)(w_L)^{1+\gamma}(w_H)^{1+\gamma}}{\pi(1-\phi)(w_H)^{1+\gamma} - \theta_H(w_L)^{1+\gamma}}\right]^{\frac{1}{\gamma}} \tag{A.14}$$

$$y_H^1 = \left[\frac{\lambda^1\phi(w_H)^{1+\gamma}}{(1-\pi)\phi + \theta_H}\right]^{\frac{1}{\gamma}} \tag{A.15}$$

Equations (3.5) and (3.6) can be combined to yield:

$$\begin{aligned} c_H^1 = \frac{1-\phi}{\alpha}[&\frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} - \delta\alpha c_H^2 + \delta\frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma} + \alpha y_L^1 \\ &+ \frac{\alpha\phi}{1-\phi}y_H^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} + \delta\alpha c_L^2 - \delta\frac{1}{1+\gamma}\left(\frac{y_L^2}{w_H}\right)^{1+\gamma}] \end{aligned} \tag{A.16}$$

Finally, (3.5) can then be used to determine $c_L^1$.

**Pooling Income Taxation: Further Details**

The relevant first-order conditions corresponding to programme (3.7) – (3.9) can be written as:

$$\pi(1-\phi)\alpha - \lambda^2(1-\phi) - \theta_H^2\alpha = 0 \tag{A.17}$$

$$-\pi(1-\phi)\left(\frac{y_L^2}{w_L}\right)^\gamma \frac{1}{w_L} + \lambda^2(1-\phi) + \theta_H^2\left(\frac{y_L^2}{w_H}\right)^\gamma \frac{1}{w_H} = 0 \tag{A.18}$$

$$(1-\pi)\phi\alpha - \lambda^2\phi + \theta_H^2\alpha = 0 \tag{A.19}$$

$$-(1-\pi)\phi\left(\frac{y_H^2}{w_H}\right)^\gamma \frac{1}{w_H} + \lambda^2\phi - \theta_H^2\left(\frac{y_H^2}{w_H}\right)^\gamma \frac{1}{w_H} = 0 \tag{A.20}$$

where $\lambda^2$ is the multiplier on the government's second-period budget constraint (3.8), and $\theta_H^2$ is the multiplier on the high-skill type's incentive-compatibility constraint (3.9). Equations (A.17) and (A.19) can be combined to yield:

$$\lambda^2 = \alpha\left[\pi(1-\phi) + (1-\pi)\phi\right] \tag{A.21}$$

$$\theta_H^2 = \frac{\lambda^2\phi}{\alpha} - (1-\pi)\phi \tag{A.22}$$

Equations (A.18) and (A.20) can be manipulated to yield, respectively:

$$y_L^2 = \left[\frac{\lambda^2(1-\phi)(w_L)^{1+\gamma}(w_H)^{1+\gamma}}{\pi(1-\phi)(w_H)^{1+\gamma} - \theta_H^2(w_L)^{1+\gamma}}\right]^{\frac{1}{\gamma}} \tag{A.23}$$

$$y_H^2 = \left[\frac{\lambda^2\phi(w_H)^{1+\gamma}}{(1-\pi)\phi + \theta_H^2}\right]^{\frac{1}{\gamma}} \tag{A.24}$$

Equations (3.8) and (3.9) can be combined to yield:

$$c_H^2 = \frac{1-\phi}{\alpha}\left[\frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma} + \alpha y_L^2 + \frac{\alpha\phi}{1-\phi}y_H^2 - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_H}\right)^{1+\gamma}\right] \tag{A.25}$$

Finally, (3.8) can then be used to determine $c_L^2$.

To solve programme (3.10) – (3.11), note that (3.11) will hold with equality at an optimum. Therefore, one can replace $c^1$ with $y^1$ in (3.10) and obtain the following first-order condition on $y^1$:

$$\pi(1-\phi)\alpha - \pi(1-\phi)\left(\frac{y^1}{w_L}\right)^\gamma \frac{1}{w_L} + (1-\pi)\phi\alpha - (1-\pi)\phi\left(\frac{y^1}{w_H}\right)^\gamma \frac{1}{w_H} = 0 \tag{A.26}$$

Using (A.26) we obtain:

$$c^1 = y^1 = \left[ \frac{\alpha \left[ \pi(1-\phi) + (1-\pi)\phi \right] (w_L)^{1+\gamma}(w_H)^{1+\gamma}}{\pi(1-\phi)(w_H)^{1+\gamma} + (1-\pi)\phi(w_L)^{1+\gamma}} \right]^{\frac{1}{\gamma}} \tag{A.27}$$

**Linear Income Taxation: Further Details**

We solve the optimal linear income tax problem for the baseline case of $\gamma = 1$ only, since the case when $\gamma \neq 1$ is highly complex.

It is straightforward to show that the solution to programme (3.12) – (3.13) yields $c_i^t = a^t + \alpha(1-\tau^t)w_i(1-\tau^t)w_i$ and $l_i^t = \alpha(1-\tau^t)w_i$. The relevant first-order conditions corresponding to programme (3.14) – (3.15) can be written as:

$$\pi(1-\phi)\alpha \frac{\partial c_L^t(\cdot)}{\partial a^t} + (1-\pi)\phi\alpha \frac{\partial c_H^t(\cdot)}{\partial a^t} - \lambda^t = 0 \tag{A.28}$$

$$\begin{aligned} &\pi(1-\phi)\alpha \frac{\partial c_L^t(\cdot)}{\partial \tau^t} - \pi(1-\phi) l_L^t(\cdot) \frac{\partial l_L^t(\cdot)}{\partial \tau^t} + (1-\pi)\phi\alpha \frac{\partial c_H^t(\cdot)}{\partial \tau^t} - (1-\pi)\phi l_H^t(\cdot) \frac{\partial l_H^t(\cdot)}{\partial \tau^t} \\ &+ \lambda^t(1-\phi)w_L l_L^t(\cdot) + \lambda^t(1-\phi)\tau^t w_L \frac{\partial l_L^t(\cdot)}{\partial \tau^t} + \lambda^t \phi w_H l_H^t(\cdot) + \lambda^t \phi \tau^t w_H \frac{\partial l_H^t(\cdot)}{\partial \tau^t} = 0 \end{aligned} \tag{A.29}$$

where $\lambda^t$ is the multiplier on the government's budget constraint (3.15). Equation (A.28) implies that:

$$\lambda^t = \alpha \left[ \pi(1-\phi) + (1-\pi)\phi \right] \tag{A.30}$$

Equation (A.29) can be simplified to yield:

$$\tau^t = \frac{(1-\phi)\alpha(w_L)^2(\alpha\pi - \lambda^t) + \phi\alpha(w_H)^2(\alpha(1-\pi) - \lambda^t)}{(1-\phi)\alpha(w_L)^2(\alpha\pi - 2\lambda^t) + \phi\alpha(w_H)^2(\alpha(1-\pi) - 2\lambda^t)} \tag{A.31}$$

Equation (A.31) can be used to determine $l_i^t$, which can then be used to determine $a^t$ using (3.15). Finally, $a^t$ and $\tau^t$ can be used to determine $c_i^t$.

# 7 Appendix B: Infinite-Horizon Model

**Separating Income Taxation: Further Details**

Since the government will solve programme (3.1) – (3.3) from period 2 onwards, the solution to this programme can be found in Appendix A. The solution to programme (4.1) – (4.3) yields the same functions for the endogenous variables as does that to programme (3.4) – (3.6), except the changed nature of the high-skill type's incentive-compatibility constraint implies that $c_H^1$ is now equal to:

$$c_H^1 = \frac{1-\phi}{\alpha}\left[\frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} + \alpha y_L^1 + \frac{\alpha\phi}{1-\phi}y_H^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} + \frac{\delta}{1-\delta}\left[u_{LF}^t(\cdot) - u_{HF}^t(\cdot)\right]\right]$$

where use has been made of the fact that $\sum_{t=2}^{\infty}\delta^{t-1}u_{iF}^t(\cdot) = \frac{\delta}{1-\delta}u_{iF}^t(\cdot)$, which follows from noting that $u_{iF}^t(\cdot)$ is the same in each period.

**Pooling Income Taxation: Further Details**

Since the government will solve programme (3.10) – (3.11) in each period, the solution to this programme can be found in Appendix A.

**Linear Income Taxation: Further Details**

Since all individuals will solve programme (3.12) – (3.13) in each period and the government will solve programme (3.14) – (3.15) in each period, the solutions to these programmes can be found in Appendix A.

# References


[1] Acemoglu, D., M. Golosov and A. Tsyvinski [2008], "Markets versus Governments", *Journal of Monetary Economics*, 55, 159-189.

[2] Acemoglu, D., M. Golosov and A. Tsyvinski [2010], "Dynamic Mirrlees Taxation under Political Economy Constraints", *Review of Economic Studies*, 77, 841-881.

[3] Albanesi, S. and C. Sleet [2006], "Dynamic Optimal Taxation with Private Information", *Review of Economic Studies*, 73, 1-30.

[4] Apps, P. and R. Rees [2006], "Repeated Optimal Nonlinear Income Taxation", *mimeo.*

[5] Berliant, M. and J. Ledyard [2005], "Optimal Dynamic Nonlinear Income Taxes with No Commitment", *mimeo.*

[6] Bisin, A. and A. Rampini [2006], "Markets as Beneficial Constraints on the Government", *Journal of Public Economics*, 90, 601-629.

[7] Brett, C. and J. Weymark [2008], "Optimal Nonlinear Taxation of Income and Savings without Commitment", *Vanderbilt University Working Paper*, No. 08-W05.

[8] Dillen, M. and M. Lundholm [1996], "Dynamic Income Taxation, Redistribution, and the Rachet Effect", *Journal of Public Economics*, 59, 69-93.

[9] Fang, H. [2006], "Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices", *International Economic Review*, 47, 1151-1185.

[10] Farhi, E. and I. Werning [2008], "The Political Economy of Nonlinear Capital Taxation", *mimeo.*

[11] Gaube, T. [2007], "Optimum Taxation of Each Year's Income", *Journal of Public Economic Theory*, 9, 127-150.

[12] Goldin, C. and L. Katz [2007], "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890-2005", *NBER Working Paper*, No. 12984.

[13] Golosov, M., M. Troshkin and A. Tsyvinski [2010], "Optimal Taxation: Merging Micro and Macro Approaches", *mimeo.*

[14] Golosov, M., A. Tsyvinski and I. Werning [2006], "New Dynamic Public Finance: A User's Guide", *NBER Macroeconomics Annual*, 2006, 317-387.

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

[16] Guo, J-T. and A. Krause [2010], "Optimal Nonlinear Income Taxation with Habit Formation", *Journal of Public Economic Theory*, forthcoming.

[17] Kocherlakota, N. [2005], "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation", *Econometrica*, 73, 1587-1621.

[18] Kocherlakota, N. [2010], *The New Dynamic Public Finance*, Princeton University Press.

[19] Krause, A. [2009], "Optimal Nonlinear Income Taxation with Learning-by-Doing", *Journal of Public Economics*, 93, 1098-1110.

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

[21] Roberts, K. [1984], "The Theoretical Limits to Redistribution", *Review of Economic Studies*, 51, 177-195.

[22] Stiglitz, J. [1982], "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics*, 17, 213-240.

[23] Walker, I. and Y. Zhu [2008], "The College Wage Premium and the Expansion of Higher Education in the UK", *Scandinavian Journal of Economics*, 110, 695-709.

[24] Werning, I. [2007], "Optimal Fiscal Policy with Redistribution", *Quarterly Journal of Economics*, 122, 925-967.

**Table 1**

Baseline Parameter Values for Numerical Simulations: Two-period Model*

| | | | | | |
|---|---|---|---|---|---|
| $\pi$ | 0.55 | $\gamma$ | 1.00 | $w_L$ | 1.00 |
| $\phi$ | 0.25 | $\delta$ | 0.46 | $w_H$ | 1.60 |
| $\alpha$ | 1.00 | $r$ | 0.04 | | |

| | *Separating* | *Pooling* | *Linear* |
|---|---|---|---|
| First-period social welfare | 0.343 | 0.302 | 0.351 |
| Second-period social welfare | 0.389 | 0.359 | 0.351 |
| Discounted total | 0.521 | 0.466 | 0.511 |

* Each period is assumed to be 20 years in length.

**Figure 1**

Social Welfare under each Tax System: Two-period Model

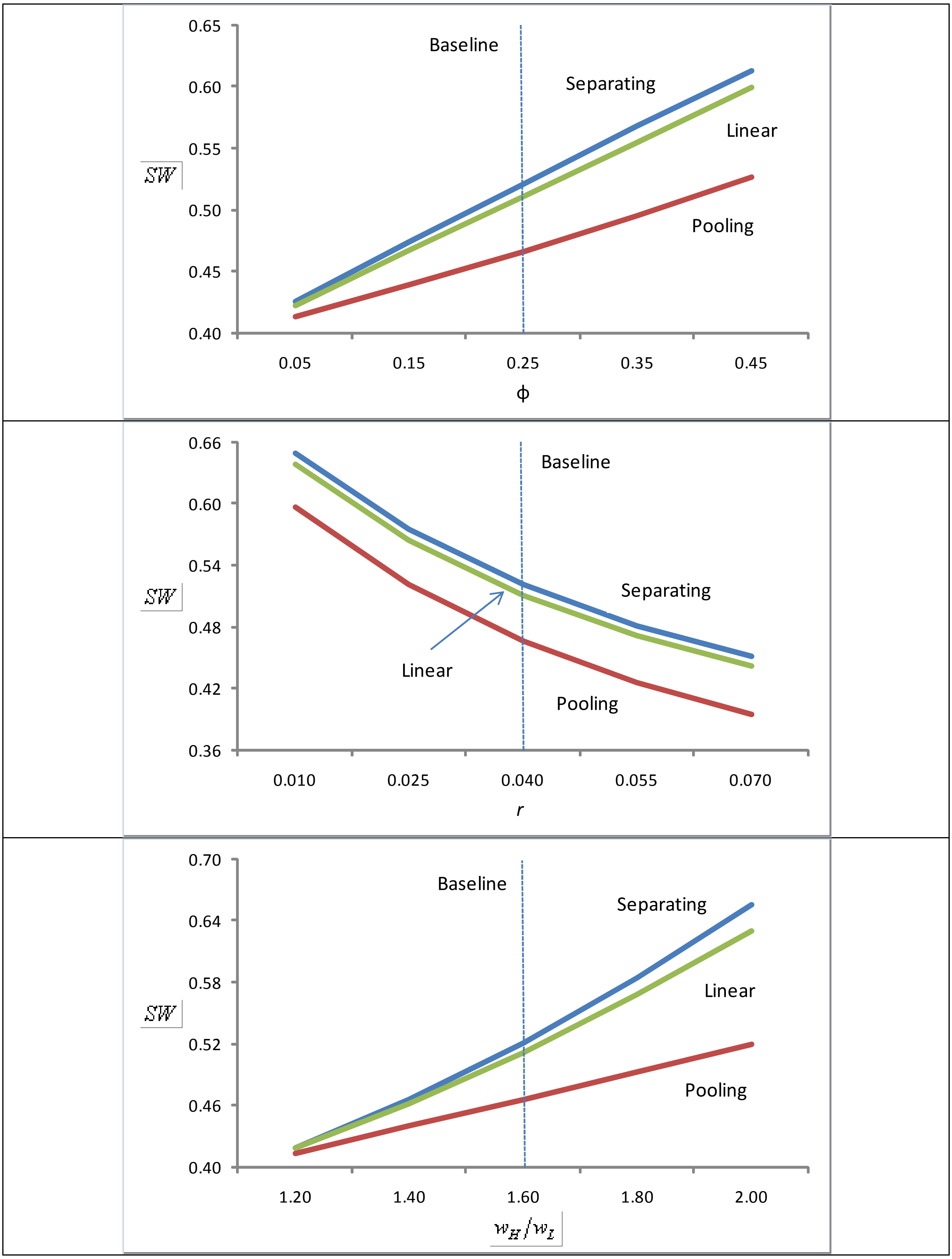

**Table 2**

Baseline Parameter Values for Numerical Simulations: Infinite-horizon Model*

| | | | | | |
|---|---|---|---|---|---|
| $\pi$ | 0.55 | $\gamma$ | 1.00 | $w_L$ | 1.00 |
| $\phi$ | 0.25 | $\delta$ | 0.96 | $w_H$ | 1.60 |
| $\alpha$ | 1.00 | $r$ | 0.04 | | |

| | *Separating* | *Pooling* | *Linear* |
|---|---|---|---|
| Social welfare | not feasible | 7.850 | 9.121 |

* Each period is assumed to be one year in length.

**Figure 2**

Social Welfare under Linear Income Taxation and Pooling: Infinite-horizon Model