Submission Number: PET11-11-00033

Dynamic income taxation without commitment: comparing alternative tax systems

Jang-ting Guo
University of California, Riverside

Alan Krause
University of York

Abstract

This paper addresses the question as to whether it is optimal to use separating or pooling nonlinear income taxation, or whether autarky is preferred, 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 taxation is optimal in the two-period model, whereas autarky is optimal when the time horizon is infinite. The effects of varying the discount rate, the degree of wage inequality, and the population of high-skill workers are also explored. For reasonable changes in these parameters, separating taxation remains optimal in the two-period model, while autarky remains optimal in the infinite-horizon model. Our results highlight the difficulty of implementing redistributive taxation when the government cannot commit.

Submitted: February 14, 2011.

Dynamic Income Taxation without Commitment: Comparing Alternative Tax Systems*

Jang-Ting Guo†
University of California, Riverside

Alan Krause‡
University of York

4 January 2011

Abstract

This paper addresses the question as to whether it is optimal to use separating or pooling nonlinear income taxation, or whether autarky is preferred, 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 taxation is optimal in the two-period model, whereas autarky is optimal when the time horizon is infinite. The effects of varying the discount rate, the degree of wage inequality, and the population of high-skill workers are also explored. For reasonable changes in these parameters, separating taxation remains optimal in the two-period model, while autarky remains optimal in the infinite-horizon model. Our results highlight the difficulty of implementing redistributive taxation when the government cannot commit.

Keywords: Dynamic Income Taxation; Commitment.

JEL Classifications: H21, H24.

---

*We thank seminar participants at the University of Manchester for many helpful comments. Any remaining errors are our responsibility.

†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)1904-423-572, Fax: 44-(0)1904-423-759, E-mail: alan.krause@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 microfoundations 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-sum taxation

[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.

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 binding 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, the autarkic equilibrium of the economy may be preferred to both separating and pooling income taxation. Little is known as to under what conditions separating taxation, pooling taxation, or autarky 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 taxation, pooling taxation, or autarky is most desirable.[6] 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.[7] We further assume that preferences take the analytically convenient additively-separable form, which allows us to conduct numerical simulations. Our main results can be summarised as follows. For empirically plausible values of the model's parameters, separating taxation is optimal in the two-period model, whereas autarky is optimal in the infinite-horizon model. We then examine how the relative desirability of separating taxation, pooling taxation, and autarky 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 taxation remains optimal in the two-period model, while autarky remains optimal in the infinite-horizon model. Pooling is not optimal in either the two-period or infinite-horizon models for all parameter changes considered. Separating 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. Autarky 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 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 autarky, separating

[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.

[6] There are other, albeit more complex, nonlinear income tax systems that the government could implement. For example, there could be partial pooling or randomised taxation. However, we restrict attention to the "pure strategy" policies of complete separating or pooling taxation. Consideration of other tax systems is left for future research.

[7] Since much of the related literature has considered either two-period or infinite-horizon settings, we also focus on these time horizons.

taxation, and pooling taxation in the two-period model, and then discusses the results of our numerical simulations. Section 4 describes the structure of autarky and each tax system when the model is extended to an infinite horizon, and then discusses the corresponding numerical simulations. Section 5 contains some concluding comments.

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 $w_H > w_L$. In order to isolate the effects of the (possible) revelation and use of skill-type information, wages are assumed to remain constant and there are no savings by individuals or the government.

Both types have the same preferences over consumption and labour in each period, which are represented by the additively-separable utility function:

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

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 $\sigma > 0$ and $\gamma > 0$ are preference parameters. When $\sigma = 1$, the utility function becomes:

$$\ln(c_i^t) - \frac{1}{1+\gamma}(l_i^t)^{1+\gamma} \tag{2.2}$$

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, where $y_i^t = w_i l_i^t$. Since there are no savings, $y_i^t - c_i^t$ is equal to total taxes paid (or, if negative, 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. Since the social welfare function is strictly concave in consumption and leisure, the government will be using its taxation powers to redistribute from high-skill 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 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 may be higher in the autarkic equilibrium, i.e., no government intervention may be optimal.

3 Two-Period Model

Deciding whether it is optimal for the government to use a nonlinear income tax system that: (i) separates in period 1 and uses first-best taxation in period 2, or (ii) pools in period 1 and uses second-best taxation in period 2, requires a comparison of social welfare in each case, and such comparisons generally depend upon the model's parameters. Accordingly, in this section we describe the structure of separating and pooling taxation, as well as the autarkic solution, which then form the basis for social-welfare comparisons made via numerical simulations.

3.1 Autarky in Both Periods

If the government does not intervene in the economy, each individual i will solve the following problem in each period. Choose c_i^t and l_i^t to maximise:

$$\frac{1}{1-\sigma}(c_i^t)^{1-\sigma} - \frac{1}{1+\gamma}\left(l_i^t\right)^{1+\gamma} \tag{3.1}$$

subject to their period t budget constraint:

$$c_i^t \leq w_i l_i^t \tag{3.2}$$

The solution to programme (3.1) – (3.2) will yield the functions $c_i^t(\sigma, \gamma, w_i)$ and $l_i^t(\cdot)$. Substituting these functions into (3.1) yields each type's utility in each period, which can then be used to determine social welfare, which we denote by $W_A^t(\phi, \sigma, \gamma, w_L, w_H)$. Total social welfare in the autarkic equilibrium is equal to $W_A^1(\cdot) + \delta W_A^2(\cdot)$.

3.2 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:

$$(1-\phi)\left[\frac{1}{1-\sigma}(c_L^2)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_L}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}(c_H^2)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma}\right] \tag{3.3}$$

subject to:

$$(1-\phi)(y_L^2 - c_L^2) + \phi(y_H^2 - c_H^2) \geq 0 \tag{3.4}$$

where (3.3) is the second-period utilitarian social welfare function, with the utility functions written in terms of the government's choice variables, while (3.4) is the government's second-period budget constraint. The solution to programme (3.3) – (3.4) yields functions for the choice variables $c_L^2(\phi, \sigma, \gamma, w_L, w_H)$, $y_L^2(\cdot)$, $c_H^2(\cdot)$ and $y_H^2(\cdot)$. Substituting these functions into (3.3) 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 (3.3) – (3.4) 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:

$$(1-\phi)\left[\frac{1}{1-\sigma}(c_L^1)^{1-\sigma}-\frac{1}{1+\gamma}\left(\frac{y_L^1}{w_L}\right)^{1+\gamma}\right]+\phi\left[\frac{1}{1-\sigma}(c_H^1)^{1-\sigma}-\frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma}\right] \quad (3.5)$$

subject to:

$$(1-\phi)(y_L^1-c_L^1)+\phi(y_H^1-c_H^1)\geq 0 \quad (3.6)$$

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

$$\frac{1}{1-\sigma}(c_L^1)^{1-\sigma}-\frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma}+\delta\left[\frac{1}{1-\sigma}(c_L^2(\cdot))^{1-\sigma}-\frac{1}{1+\gamma}\left(\frac{y_L^2(\cdot)}{w_H}\right)^{1+\gamma}\right] \quad (3.7)$$

where (3.5) is first-period social welfare, (3.6) is the government's first-period budget constraint, and (3.7) 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 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.[8] 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

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

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 (3.5)−(3.7) yields functions for the choice variables $c_L^1(\phi, \sigma, \gamma, w_L, w_H, \delta)$, $y_L^1(\cdot)$, $c_H^1(\cdot)$ and $y_H^1(\cdot)$. Substituting these functions into (3.5) yields the level of social welfare in period 1 under second-best taxation, which we denote by $W_S^1(\cdot)$. Total social welfare with first-period separation is equal to $W_S^1(\cdot) + \delta W_F^2(\cdot)$.

3.3 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:

$$(1-\phi)\left[\frac{1}{1-\sigma}(c_L^2)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_L}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}(c_H^2)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma}\right] \quad (3.8)$$

subject to:

$$(1-\phi)(y_L^2 - c_L^2) + \phi(y_H^2 - c_H^2) \geq 0 \quad (3.9)$$

$$\frac{1}{1-\sigma}(c_H^2)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^2}{w_H}\right)^{1+\gamma} \geq \frac{1}{1-\sigma}(c_L^2)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^2}{w_H}\right)^{1+\gamma} \quad (3.10)$$

where (3.8) is second-period social welfare, (3.9) is the government's second-period budget constraint, and (3.10) is the high-skill type's incentive-compatibility constraint. The solution to programme (3.8)−(3.10) yields functions for the choice variables $c_L^2(\phi, \sigma, \gamma, w_L, w_H)$, $y_L^2(\cdot)$, $c_H^2(\cdot)$ and $y_H^2(\cdot)$. Substituting these functions into (3.8) 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:

$$(1-\phi)\left[\frac{1}{1-\sigma}(c^1)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y^1}{w_L}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}(c^1)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y^1}{w_H}\right)^{1+\gamma}\right] \quad (3.11)$$

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

$$y^1 - c^1 \geq 0 \tag{3.12}$$

Since the budget constraint will bind, the solution to programme (3.11) – (3.12) will involve $c^1 = y^1 = y^1(\phi, \sigma, \gamma, w_L, w_H)$. Substituting this function into (3.11) yields the level of social welfare in period 1 under pooling, which we denote by $W_P^1(\cdot)$. Total social welfare with first-period pooling is equal to $W_P^1(\cdot) + \delta W_S^2(\cdot)$.

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 separating, pooling, and autarkic solutions. 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 OECD [2010] reports that on average across OECD countries, approximately one-quarter of all adults have attained tertiary level education. We therefore assume that 25% of individuals are high-skill workers, i.e., we set $\phi = 0.25$. The preference parameter σ is set to unity, as is common in the macroeconomics literature, while setting $\gamma = 2$ corresponds to a labour supply elasticity of 0.5,[9] which is consistent with empirical estimates. We assume an annual discount rate of 5%, which is in line with common practice. Since most individuals work for around 40 years of their lives, we take each period to be 20 years in length.[10] An annual discount rate of 5% then corresponds to a 20-year discount factor of $\delta = 0.38$. 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 taxation is social-welfare

[9] To see this, note that the first-order conditions corresponding to programme (3.1) – (3.2) can be manipulated to yield $(c_i^t)^\sigma (l_i^t)^\gamma = w_i$ or $\sigma \ln(c_i^t) + \gamma \ln(l_i^t) = \ln(w_i)$, which implies a labour supply elasticity of $1/\gamma$.

[10] Kocherlakota [2010] also assumes that each period consists of twenty years when individuals work for two periods.

maximising, autarky is ranked second, while pooling is third. Pooling is worse than autarky, even though pooling in period 1 allows second-best taxation to be used in period 2 (which is better than autarky). However, pooling in the first period is very costly, as reflected in the low level of social welfare.

Figure 1 shows the effects of relatively large variations in the size of the high-skill population (ϕ), the discount rate (r), and the wage premium (w_H), whilst holding all other parameters at their baseline levels. The social-welfare ranking of separating taxation, autarky, and pooling remains unchanged for the variations considered.[11] Separation increases its advantage over autarky and pooling as ϕ increases. An increase in ϕ 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. Autarky also increases its advantage over pooling as ϕ increases. Increases in ϕ 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 effective in achieving redistribution. But the benefit is not sufficient to overcome the increased inefficiency of pooling in the first period.

Higher values of r increase the advantage that separation has over autarky and pooling. 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 autarky more attractive than pooling. Since pooling in period 1 is less desirable than autarky, but nonlinear income taxation in period 2

[11] We also examined the effects of varying the preference parameters. Specifically, we varied σ between 0.5 and 2, and γ between 1 and 4. For these variations, the ranking of separation, autarky, and pooling remains intact. Further details are available upon request.

is better than autarky, 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.

Separation increases its advantage over autarky and pooling as w_H increases. Given the government's redistributive concerns, autarky 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 rises. As w_H increases, autarky is also increasingly preferred to pooling. On the one hand, an increase in wage inequality exacerbates the inefficiency of pooling in period 1, but on the other hand higher wage inequality increases the desirability of using nonlinear income taxation in period 2. However, on balance our numerical simulations indicate that pooling becomes increasingly less desirable than autarky as w_H increases.

4 Infinite-Horizon Model

In this section, we describe how the general structure of autarky, separating taxation, and pooling taxation changes when the model is extended from two periods to an infinite-horizon setting.

4.1 Autarky in Each Period

If the government does not intervene, individuals will solve programme $(3.1) - (3.2)$ in each period. Total social welfare under autarky is therefore equal to $\sum\limits_{t=1}^{\infty} \delta^{t-1} W_A^t(\cdot)$.

4.2 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 $(3.3) - (3.4)$ in periods $2, ..., \infty$. Let $u_{iF}^t(\phi, \sigma, \gamma, w_L, w_H)$ denote the utility type i obtains under first-best taxation in each period, let $\widehat{u}_{HF}^t(\cdot)$ denote the utility the high-skill type obtains from the low-skill type's first-best tax treatment in each period, and let $W_F^t(\cdot)$ 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 (3.3) – (3.4) in periods $2, ..., \infty$. Therefore, in order to induce individuals to reveal their types 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 types, respectively, to maximise:

$$(1-\phi)\left[\frac{1}{1-\sigma}(c_L^1)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_L}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}(c_H^1)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma}\right] \quad (4.1)$$

subject to:

$$(1-\phi)(y_L^1 - c_L^1) + \phi(y_H^1 - c_H^1) \geq 0 \quad (4.2)$$

$$\frac{1}{1-\sigma}(c_H^1)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} + \sum_{t=2}^{\infty}\delta^{t-1}u_{HF}^t(\cdot) \geq \frac{1}{1-\sigma}(c_L^1)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} + \sum_{t=2}^{\infty}\delta^{t-1}\widehat{u}_{HF}^t(\cdot) \quad (4.3)$$

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 $c_L^1(\phi, \sigma, \gamma, w_L, w_H, \delta)$, $y_L^1(\cdot)$, $c_H^1(\cdot)$ 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 with separation is equal to $W_S^1(\cdot) + \sum_{t=2}^{\infty}\delta^{t-1}W_F^t(\cdot)$.

4.3 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 (3.11) − (3.12) in each period. Total social welfare under pooling is therefore equal to $\sum_{t=1}^{\infty} \delta^{t-1} W_P^t(\phi, \sigma, \gamma, w_L, w_H)$, where $W_P^t(\cdot)$ is the level of social welfare associated with programme (3.11) − (3.12).

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 5% corresponds to a one-year discount factor of $\delta = 0.95$. For the baseline parameter values, separating 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 9%, which seems unlikely. Therefore, the only options available to the government in the infinite-horizon model are to pool the individuals in every period or to not intervene, i.e., allow the autarkic equilibrium to be realised. Since pooling is extreme in that it imposes the same consumption/pre-tax income allocation on both types, the autarkic solution is better in the infinite-horizon model.

Figure 2 shows the effects of varying the high-skill population ϕ, the discount rate r, and the wage premium w_H on the relative desirability of autarky and pooling. Autarky remains preferred for all variations considered.[12] For increases in ϕ and w_H, autarky 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 autarky and pooling appears to narrow, but proportionally the advantage autarky has over pooling remains constant. This is because the same

[12] We also again examined the effects of varying σ between 0.5 and 2, and γ between 1 and 4. For these variations, separation remains infeasible, and autarky maintains its advantage over pooling. Further details are available upon request.

allocation is implemented under autarky 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. When the discount rate reaches 9%, separating taxation becomes feasible, but the associated level of social welfare is lower than both autarky and pooling. Hence separating taxation becomes feasible, but it would never be chosen.

5 Concluding Comments

This paper has addressed the question as to whether it is optimal to use separating or pooling nonlinear income taxation, or whether autarky is preferred, 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 taxation is optimal in the two-period model, whereas autarky 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 or autarky is optimal. We have also examined how the relative desirability of each tax system and autarky is affected by changes in the model's parameters. In light of the fact that autarky can be the best option, our results highlight the difficulty of implementing redistributive taxation when the government cannot commit to its future tax policy.

Deciding whether separating taxation, pooling taxation, or autarky is most desirable requires a comparison of social welfare in each case. In order to make such comparisons, we have used a simple model that lends itself to numerical simulations. The question remains as to how dependent our results are on the specifics of the model. Based on the nature of the intuition driving our results, we conjecture that our main conclusions would hold-up in more general settings. Nevertheless, a number of extensions of our work seem worth pursuing, although some are likely to be more worthwhile than others. One could extend the model to more than two skill types, but there will still be one group of individuals (the higher skilled) who want to conceal their type, and another group (the lower skilled) who want type information revealed. Accordingly, our main results are unlikely to be affected by considering more types. An extension that seems more interesting would be to allow for savings by individuals and the government. Allowing the government to transfer resources over time could help it overcome some of the inefficiencies associated with separating and pooling taxation. However, such efforts by the government might be undermined by individual savings behaviour. Another possible extension is to examine the "mixed strategy" policies of partial pooling or randomised taxation. These seem interesting avenues for future research.

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

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

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

[11] 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.

[12] Golosov, M., M. Troshkin and A. Tsyvinski [2010], "Optimal Taxation: Merging Micro and Macro Approaches", *Journal of Money, Credit and Banking*, forthcoming.

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

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

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

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

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

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

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

[20] OECD [2010], *Highlights from Education at a Glance*, OECD publications 2010.

[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] 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*

ϕ	0.25	δ	0.38	w_L	1.00
σ	1.00	r	0.05	w_H	1.60
γ	2.00				

	Separating	*Pooling*	*Autarky*
First-period social welfare	−0.210	−0.264	−0.216
Second-period social welfare	−0.181	−0.196	−0.216
Discounted total	−0.279	−0.337	−0.297

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

FIGURE 1

Social Welfare: Two-period Model

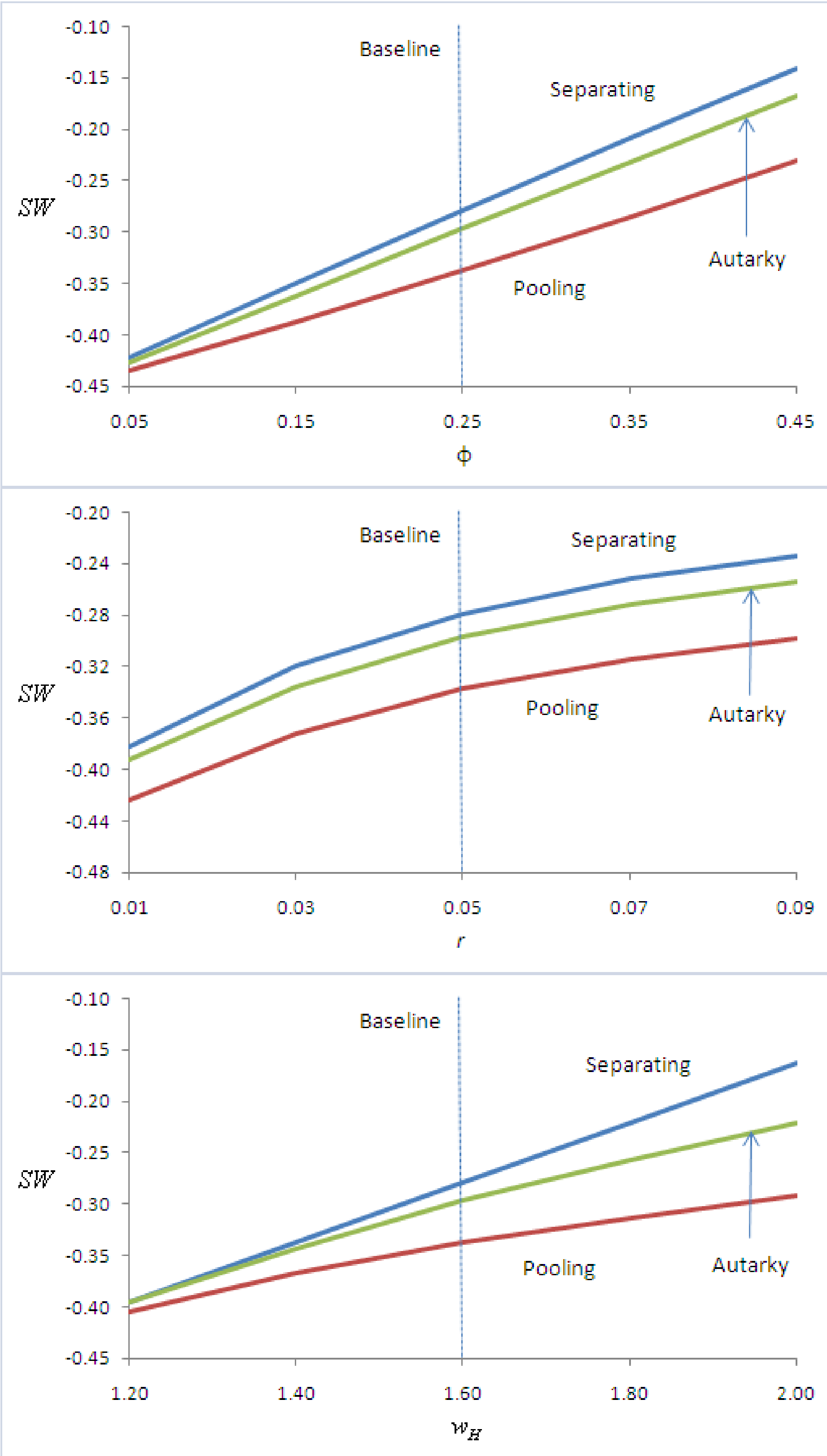

TABLE 2

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

ϕ	0.25	δ	0.95	w_L	1.00
σ	1.00	r	0.05	w_H	1.60
γ	2.00				

	Separating	*Pooling*	*Autarky*
Social welfare	not feasible	−5.534	−4.532

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

FIGURE 2

Social Welfare: Infinite-horizon Model