Macroeconomic Dynamics, **18**, 2014, 1403–1427. Printed in the United States of America.
doi:10.1017/S1365100512001010

OPTIMAL DYNAMIC NONLINEAR INCOME TAXATION UNDER LOOSE COMMITMENT

JANG-TING GUO
University of California, Riverside

ALAN KRAUSE
University of York

This paper examines an infinite-horizon model of nonlinear income taxation in which the probability that the government can commit is high, but not certain. In this “loose commitment” environment, we find that even a little uncertainty over whether the government can commit yields substantial effects on the optimal dynamic nonlinear income tax system. This result holds even though separating taxation remains optimal, as in the case of full commitment. Under an empirically plausible parameterization, our numerical simulations show that high-skilled individuals must be subsidized in the short run, despite the government’s redistributive objective, unless the probability of commitment is higher than 98%. Loose commitment also reverses the short-run welfare effects of changes in most of the model’s parameters, and yields some counterintuitive outcomes. For example, all individuals are worse off, rather than better off, in the short run when the proportion of high-skilled individuals in the economy increases.

Keywords: Dynamic Nonlinear Income Taxation, Loose Commitment

1. INTRODUCTION

There is currently a great deal of interest in dynamic nonlinear income taxation, as exemplified by the “new dynamic public finance” literature that extends the static Mirrlees (1971) model of nonlinear income taxation to a dynamic setting.[1] In the Mirrlees model, individuals are distinguished by their skill levels, which results in differences in their income-earning abilities. However, the government cannot implement (the first-best) personalized lump-sum taxation based on skills as the Second Welfare Theorem would recommend, owing to the assumption that each individual’s skill type is private information. Instead, the government can only implement (the second-best) incentive-compatible nonlinear income taxation, under which each individual is willing to reveal his or her skill type. In dynamic

We thank an Associate Editor and two anonymous referees for very helpful comments and suggestions that improved the paper. We have also benefitted from discussions with Fabrizio Iacone and Makoto Shimoji regarding Markov switching processes. Any remaining errors are our responsibility. Address correspondence to: Jang-Ting Guo, Department of Economics, University of California, Riverside, Riverside, CA 92521-0427, USA; e-mail: guojt@ucr.edu.

 1365-1005/13

versions of the Mirrlees model, however, skill-type information revealed in period 1 could, in principle, be used by the government to implement personalized lump-sum taxation from period 2 onward. This feature makes period 1 somewhat special in dynamic Mirrlees models. For the sake of analytical simplicity, the new dynamic public finance literature typically assumes that the government can commit to its future tax policy. That is, the government continues to implement incentive-compatible taxation even after individuals have revealed their types.

It seems possible to make convincing arguments in favor of assuming either commitment or no commitment. For example, one might defend the commitment assumption on the basis that real-world income tax systems are not frequently redesigned,[2] and that there are long-run benefits to be gained by a government that makes and keeps its promises. On the other hand, the commitment assumption has been criticized as being unrealistic, because the present government cannot easily impose binding constraints on the policies of future governments.[3] Accordingly, although most of the previous literature assumes full commitment, there are some works that consider the opposite case of no-commitment. Brett and Weymark (2008a) and Farhi et al. (2012) examine dynamic Mirrlees models in which the government may impose nonlinear taxes on both savings/capital and labor income. Despite some interesting differences in their models, both papers find that zero taxation of savings/capital is optimal under commitment, but savings/capital should be taxed under no commitment. These authors therefore offer a new argument in favor of savings/capital taxation based on the inability of the government to commit. Other models of dynamic nonlinear income taxation without commitment have been developed by Apps and Rees (2006), Krause (2009), and Guo and Krause (2011b, 2013), among others. Although savings do not feature in these papers, the focus is again on comparing outcomes under full commitment versus no commitment. In contrast, Battaglini and Coate (2008) examine a dynamic model of nonlinear income taxation in which individuals' high or low skill types are stochastic. Therefore, even if the government cannot commit, the advantage it obtains from acquiring skill-type information in any particular period is diminished, because an individual may change type. As a result, second-best income tax systems can be time-consistent provided the correlation in types is not perfect.

Because the assumptions of commitment or no commitment can be viewed as polar cases, we depart from the existing literature by assuming that the government can commit not to use skill-type information only with some probability. When the government cannot commit fully, however, it is well known that it may no longer be social-welfare maximizing for the government to implement (separating) nonlinear income taxation in which individuals are willing to reveal their types.[4] Instead, it may be optimal to pool the individuals—by imposing the same tax treatment on everyone—so that type information is not revealed. To avoid this possibility and ensure that separating taxation remains optimal as under full commitment, we postulate that the probability of commitment is sufficiently high; hence the term "loose commitment."[5] Specifically, loose commitment is modeled as a Markov switching process, whereby *in each period* there is some probability that the government can and cannot commit. To the best of our knowledge, this paper is the *first* to examine dynamic nonlinear income taxation in a loose commitment framework. For the purpose of isolating the taxation impact of loose commitment, each agent's skill type is postulated to be time-invariant, and there are no savings by individuals or by the government. Moreover, we consider the two-type version of the Mirrlees model introduced by Stiglitz (1982), but extend it to an infinite-horizon setting. We further assume that the utility function is additively separable between consumption and labor. These simplifications allow us to investigate in detail the effects of loose commitment.

More specifically, the aim of our paper is to investigate how changes in the probability of commitment affect optimal dynamic nonlinear income taxation. However, the literature that examines the comparative statics of optimal nonlinear income taxes in the static Mirrlees model has shown that analytical results can be obtained only when the utility function is quasi-linear.[6] We cannot assume quasi-linearity in our model, because the solution to the first-best taxation problem becomes indeterminate.[7] Moreover, our dynamic version of the Mirrlees model is significantly more complicated than its static counterpart. For these reasons, we rely on numerical simulations to illustrate the impact of loose commitment.

Our main finding is that even a small amount of uncertainty regarding whether the government can commit yields a substantial effect on optimal dynamic nonlinear income taxation. This result holds even though separating taxation remains optimal, as in the case of full commitment. Under an empirically plausible parameterization, our quantitative analysis shows that high-skilled individuals must be subsidized in period 1, despite the government's redistributive concerns, unless the probability of commitment is greater than 98%. This is because high-skilled individuals know that if they reveal their type, the government will occasionally deviate from implementing (the second-best) incentive-compatible taxation to implement (the first-best) personalized lump-sum taxation. Therefore, high-skilled individuals require compensation in period 1 if they are to reveal their type. Loose commitment also reverses the short-run welfare effects of changes in most of the model's parameters. For example, all individuals are worse off, rather than better off, in period 1 when the proportion of high-skilled individuals in the economy increases. This counterintuitive result can be understood as follows. High-skilled individuals are worse off in period 1 when their population rises because they are better off in the long run, which means that they require less compensation in period 1 to reveal their type. But low-skilled individuals are also worse off in period 1, because each low-skilled individual must pay more tax to finance the larger total subsidy received by the increased population of high-skilled individuals. The short-run welfare effects of varying the high-skilled type's wage, the discount rate, and the labor supply elasticity are also shown to be affected—and often reversed—by loose commitment.

The remainder of the paper is organized as follows. Section 2 describes the analytical framework that we consider, whereas Section 3 analyzes the structure of optimal dynamic nonlinear income taxation under loose commitment. The results of our numerical simulations are discussed in Section 4, whereas Section 5 contains some concluding comments and suggestions for future research.

2. ANALYTICAL FRAMEWORK

There is a unit measure of infinitely lived individuals, with a proportion $\phi \in (0, 1)$ being high-skilled workers and the remaining $(1 - \phi)$ being low-skilled workers. The high-skilled type's wage is denoted by $w_{\rm H}$, whereas that for the low-skilled type is denoted by $w_{\rm L}$, where $w_{\rm H} > w_{\rm L}$. To isolate the dynamic effects of loose commitment, wages are assumed to remain constant through time and there are no savings by individuals or the government. Thus the only link between periods in our model is the revelation and possible use of skill-type information. The preferences of both types of individual in each period are represented by the (analytically convenient) additively separable utility function

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

where c_i^t denotes type i's consumption in period t, l_i^t denotes type i's labor supply in period t, and $\sigma > 0$ and $\gamma > 0$ are preference parameters. When $\sigma = 1$ the utility function becomes

$$\ln\left(c_i^t\right) - \frac{1}{1+\gamma}\left(l_i^t\right)^{1+\gamma}. \tag{2}$$

Both types of individual 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 given by $y_i^t = w_i l_i^t$, and because individuals cannot save or borrow, $y_i^t - c_i^t$ is equal to taxes paid (or, if negative, transfers received) by a type i individual in period t.

The government uses its taxation powers to maximize social welfare, which is assumed to be measurable by a utilitarian social welfare function. The government will therefore have a redistributive objective, meaning it will be seeking to tax high-skilled individuals in order to subsidize low-skilled individuals. However, the government cannot implement (the first-best) personalized lump-sum taxation in every period, as each individual's skill type is initially private information. Because the government cannot commit with certainty, individuals know that once they reveal their type, they may be subjected to first-best taxation. This means that some individuals, namely, high-skilled individuals, have to be compensated if they are to be willing to reveal their type, and this compensation is potentially very costly from the government's perspective of maximizing social welfare. Accordingly, rather than design a "separating" tax system in period 1 in which individuals are willing to reveal their types, it may be optimal for the government to use "pooling" taxation for some period of time in which type information is not revealed.[8]

In general, if agents' types are separated in period T, it is assumed that all individuals know that the government will use second-best taxation in period $T+1$ with probability p, and will use first-best taxation in period $T+1$ with probability $(1-p)$. That is, commitment occurs with probability p, and no commitment occurs with probability $(1-p)$. Then, from period $T+1$ onward, the probability that the government will use second-best or first-best taxation follows a Markov switching process according to the following transition probabilities:

$$\Pr(\text{SB in period } t+1 \mid \text{SB in period } t) = q_{\text{S}}, \tag{3}$$

$$\Pr(\text{FB in period } t+1 \mid \text{FB in period } t) = q_{\text{F}}. \tag{4}$$

That is, if the government uses second-best (SB) taxation in period t (where $t \geq T+1$), there is a probability of q_{S} that it will use second-best taxation again in period $t+1$, and a probability of $(1-q_{\text{S}})$ that it will switch and use first-best (FB) taxation in period $t+1$. Likewise, if the government uses first-best taxation in period t (where $t \geq T+1$), it uses first-best taxation again in period $t+1$ with probability q_{F}, and it uses second-best taxation in period $t+1$ with probability $(1-q_{\text{F}})$.

Our specification allows full commitment and no commitment as special cases. Under full commitment, $p=1$ and $q_{\text{S}}=1$, in which case the government always uses second-best taxation. Under no commitment, $p=0$ and $q_{\text{F}}=1$, in which case the government always uses first-best taxation once skill types have been revealed. However, because we are interested in loose commitment, our analysis does not explore these polar cases. This means that the government may switch between using second-best and first-best taxation across two consecutive time periods. One could justify this formulation in a number of ways. For example, the incumbent government may keep its own promise not to use skill-type information in the next period, but a newly elected government may not feel bound by the previous government's promise. Alternatively, because low-skilled individuals are better off under first-best taxation and high-skilled individuals are better off under second-best taxation, one can imagine that a left-wing government is more likely to implement the former and a right-wing government is more likely to implement the latter. Finally, one can think of the same government remaining in power and being able to commit with a high probability, but pressure from low-skilled individuals causes it to occasionally deviate and implement first-best taxation. Because the focus of this paper is to examine the effects of relaxing the standard full-commitment assumption "a little bit," all of these possible interpretations are consistent with our main objective.

To summarize, the timing in our model is as follows:

1. At the beginning of period 1, the government knows there are $\phi \in (0,1)$ high-skilled individuals and $(1-\phi)$ low-skilled individuals in the economy, but it does not know each individual's skill type.

2. The government uses separating taxation in some period T to obtain skill-type information, and it uses pooling taxation in periods 1 to $T-1$ (if $T \geq 2$).
3. All individuals know that if the government uses separating taxation in period T, it will use second-best taxation in period $T+1$ with probability p, and it will use first-best taxation in period $T+1$ with probability $(1-p)$.
4. From period $T+1$ onward, the probability that the government uses second-best or first-best taxation in each period follows a Markov switching process according to the transition probabilities in equations (3) and (4).

3. OPTIMAL TAXATION UNDER LOOSE COMMITMENT

Our analysis begins by describing first-best and second-best taxation, which the government may use after the types have been separated. We then describe the nature of taxation up to the separation period.

3.1. First-Best Taxation

If the government uses first-best taxation in period t, it can be described as choosing tax treatments $\langle c_{\mathrm{L}}^t, y_{\mathrm{L}}^t \rangle$ and $\langle c_{\mathrm{H}}^t, y_{\mathrm{H}}^t \rangle$ for the low-skilled and high-skilled individuals, respectively, to maximize

$$(1-\phi)\left[\frac{1}{1-\sigma}\left(c_{\mathrm{L}}^t\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_{\mathrm{L}}^t}{w_{\mathrm{L}}}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}\left(c_{\mathrm{H}}^t\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_{\mathrm{H}}^t}{w_{\mathrm{H}}}\right)^{1+\gamma}\right], \quad (5)$$

subject to

$$(1-\phi)\left(y_{\mathrm{L}}^t - c_{\mathrm{L}}^t\right) + \phi\left(y_{\mathrm{H}}^t - c_{\mathrm{H}}^t\right) \geq 0, \quad (6)$$

where equation (5) is the utilitarian social welfare function, and equation (6) is the government's budget constraint.[9] As the government knows each individual's type and is using this information, low-skilled individuals must accept $\langle c_{\mathrm{L}}^t, y_{\mathrm{L}}^t \rangle$ and high-skilled individuals must accept $\langle c_{\mathrm{H}}^t, y_{\mathrm{H}}^t \rangle$. That is, there is no private information, so individuals cannot deviate from their intended tax treatments.

The solution to program (5)–(6) yields $c_{\mathrm{L}}^t(\phi, \sigma, \gamma, w_{\mathrm{L}}, w_{\mathrm{H}})$, $y_{\mathrm{L}}^t(\cdot)$, $c_{\mathrm{H}}^t(\cdot)$, and $y_{\mathrm{H}}^t(\cdot)$. Substituting these functions into the utility function yields $u_{iF}^t(\cdot)$, which denotes the utility a type i individual obtains under first-best taxation in period t. Likewise, let $W_{\mathrm{FB}}^t(\cdot)$ denote the level of social welfare under first-best taxation in period t.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

3.2. Second-Best Taxation

If the government uses second-best taxation in period t, it can be described as choosing tax treatments $\langle c_L^t, y_L^t \rangle$ and $\langle c_H^t, y_H^t \rangle$ for the low-skilled and high-skilled individuals, respectively, to maximize

$$(1-\phi)\left[\frac{1}{1-\sigma}\left(c_L^t\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^t}{w_L}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}\left(c_H^t\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^t}{w_H}\right)^{1+\gamma}\right], \quad (7)$$

subject to

$$(1-\phi)\left(y_L^t - c_L^t\right) + \phi\left(y_H^t - c_H^t\right) \geq 0, \quad (8)$$

$$\frac{1}{1-\sigma}\left(c_H^t\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_H^t}{w_H}\right)^{1+\gamma} \geq \frac{1}{1-\sigma}\left(c_L^t\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_L^t}{w_H}\right)^{1+\gamma}, \quad (9)$$

where equation (7) is the utilitarian social welfare function, equation (8) is the government's budget constraint, and equation (9) is the high-skilled type's incentive-compatibility constraint.[10] Even though the government knows each individual's skill type, and therefore has enough information to implement first-best taxation, commitment implies that the government does not use this information. It therefore implements incentive-compatible taxation. Note that if the government could commit with certainty, it would simply solve program (7)–(9) in each period.

The solution to program (7)–(9) yields the functions $c_L^t(\phi, \sigma, \gamma, w_L, w_H)$, $y_L^t(\cdot)$, $c_H^t(\cdot)$, and $y_H^t(\cdot)$. Substituting these functions into the utility function yields $u_{iS}^t(\cdot)$, which denotes the utility a type i individual obtains under second-best taxation in period t. Likewise, let $W_{SB}^t(\cdot)$ denote the level of social welfare under second-best taxation in period t.

3.3. Optimal Taxation in the Separating Period

Suppose the government implements separating taxation in period T. From period $T+1$ onward, the "continuation utility" of a type i individual can be written as the following recursive equations:

$$V_{iS}^t = u_{iS}^t(\cdot) + q_S \delta V_{iS}^{t+1} + (1-q_S)\delta V_{iF}^{t+1}, \quad (10)$$

$$V_{iF}^t = u_{iF}^t(\cdot) + q_F \delta V_{iF}^{t+1} + (1-q_F)\delta V_{iS}^{t+1}, \quad (11)$$

where V_{iS}^t (resp. V_{iF}^t) is type i's continuation utility if second-best (resp. first-best) taxation is used in period t (where $t \geq T+1$). For example, equation (10) can be interpreted as follows. If the government uses second-best taxation in period t, type i individual obtains his or her second-best utility levels $u_{iS}^t(\cdot)$ in period t, and with probability q_S they continue to obtain their second-best utility levels in period $t+1$, but with probability $(1-q_S)$ the government switches and they obtain their first-best utility levels in period $t+1$. That is, the continuation utility function V_{iF}^{t+1} becomes active. Similarly, the continuation utility of a mimicking high-skilled individual can be written as

$$V_{\mathrm{MS}}^{t} = u_{HS}^{t}(\cdot) + q_{\mathrm{S}}\delta V_{\mathrm{MS}}^{t+1} + (1-q_{\mathrm{S}})\delta V_{\mathrm{MF}}^{t+1}, \tag{12}$$

$$V_{\mathrm{MF}}^{t} = u_{\mathrm{MF}}^{t}(\cdot) + q_{\mathrm{F}}\delta V_{\mathrm{MF}}^{t+1} + (1-q_{\mathrm{F}})\delta V_{\mathrm{MS}}^{t+1}, \tag{13}$$

where $u_{\mathrm{MF}}^{t}(\cdot)$ denotes the utility a high-skilled individual obtains in period t from the low-skilled type's first-best tax treatment.

The government's behavior in the separation period can now be described as follows. Choose tax treatments $\langle c_{\mathrm{L}}^{T}, y_{\mathrm{L}}^{T}\rangle$ and $\langle c_{\mathrm{H}}^{T}, y_{\mathrm{H}}^{T}\rangle$ for the low-skilled and high-skilled individuals, respectively, to maximize

$$(1-\phi)\left[\frac{1}{1-\sigma}\left(c_{\mathrm{L}}^{T}\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_{\mathrm{L}}^{T}}{w_{\mathrm{L}}}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}\left(c_{\mathrm{H}}^{T}\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_{\mathrm{H}}^{T}}{w_{\mathrm{H}}}\right)^{1+\gamma}\right], \tag{14}$$

subject to

$$(1-\phi)\left(y_{\mathrm{L}}^{T} - c_{\mathrm{L}}^{T}\right) + \phi\left(y_{\mathrm{H}}^{T} - c_{\mathrm{H}}^{T}\right) \geq 0, \tag{15}$$

$$\frac{1}{1-\sigma}\left(c_{\mathrm{H}}^{T}\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_{\mathrm{H}}^{T}}{w_{\mathrm{H}}}\right)^{1+\gamma} + p\delta V_{HS}^{T+1}(\cdot) + (1-p)\delta V_{HF}^{T+1}(\cdot) \geq \frac{1}{1-\sigma}\left(c_{\mathrm{L}}^{T}\right)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y_{\mathrm{L}}^{T}}{w_{\mathrm{H}}}\right)^{1+\gamma} + p\delta V_{\mathrm{MS}}^{T+1}(\cdot) + (1-p)\delta V_{\mathrm{MF}}^{T+1}(\cdot), \tag{16}$$

where equation (14) is the utilitarian social welfare function, equation (15) is the government's budget constraint, and equation (16) is the high-skilled type's incentive-compatibility constraint.[11] High-skilled individuals who reveal their type in period T by choosing $\langle c_{\mathrm{H}}^{T}, y_{\mathrm{H}}^{T}\rangle$ can expect to obtain a continuation utility from period $T+1$ of $V_{\mathrm{HS}}^{T+1}(\cdot)$ with probability p, and $V_{\mathrm{HF}}^{T+1}(\cdot)$ with probability $(1-p)$. On the other hand, high-skilled individuals who mimic in period T by choosing $\langle c_{\mathrm{L}}^{T}, y_{\mathrm{L}}^{T}\rangle$ can expect to obtain a continuation utility from period $T+1$ of $V_{\mathrm{MS}}^{T+1}(\cdot)$ with probability p, and $V_{\mathrm{MF}}^{T+1}(\cdot)$ with probability $(1-p)$. Therefore, in order to induce high-skilled individuals to reveal their type in period T, the tax treatments must satisfy equation (16).

The solution to program (14)–(16) yields $c_{\mathrm{L}}^{T}(\phi, \sigma, \gamma, w_{\mathrm{L}}, w_{\mathrm{H}}, \delta, p, q_{\mathrm{S}}, q_{\mathrm{F}})$, $y_{\mathrm{L}}^{T}(\cdot)$, $c_{\mathrm{H}}^{T}(\cdot)$, and $y_{\mathrm{H}}^{T}(\cdot)$. Substituting these functions into the utility function yields $u_{i\mathrm{Sep}}^{T}(\cdot)$, which denotes the utility a type i individual obtains in the separation period. Likewise, let $W_{\mathrm{Sep}}^{T}(\cdot)$ denote the level of social welfare in the separation period.

3.4. Optimal Taxation before the Separating Period

If the government chooses to use separating taxation in period 1, then aggregate social welfare over the infinite time horizon is equal to

$$\mathrm{SW}_{\mathrm{Sep}}^{1} = W_{\mathrm{Sep}}^{1}(\cdot) + p\delta Z_{\mathrm{SB}}^{2} + (1-p)\delta Z_{\mathrm{FB}}^{2}, \quad (17)$$

where

$$Z_{\mathrm{SB}}^{T+1} = W_{\mathrm{SB}}^{T+1}(\cdot) + q_{\mathrm{S}}\delta Z_{\mathrm{SB}}^{T+2} + (1-q_{\mathrm{S}})\delta Z_{\mathrm{FB}}^{T+2}, \quad (18)$$

$$Z_{\mathrm{FB}}^{T+1} = W_{\mathrm{FB}}^{T+1}(\cdot) + q_{\mathrm{F}}\delta Z_{\mathrm{FB}}^{T+2} + (1-q_{\mathrm{F}})\delta Z_{\mathrm{SB}}^{T+2}, \quad (19)$$

where Z_{SB}^{T+1} (resp. Z_{FB}^{T+1}) denotes continuation social welfare if second-best (resp. first-best) taxation is used in period $T+1$.

However, it may be optimal for the government to pool the individuals for $T-1$ periods before using separating taxation in period T. In this case, aggregate social welfare over the infinite time horizon is equal to

$$\mathrm{SW}_{\mathrm{Sep}}^{T} = \sum_{t=1}^{T-1} \delta^{t-1} W_{\mathrm{Pool}}^{t}(\cdot) + \delta^{T-1} W_{\mathrm{Sep}}^{T}(\cdot) + p\delta^{T} Z_{\mathrm{SB}}^{T+1} + (1-p)\delta^{T} Z_{\mathrm{FB}}^{T+1}, \quad (20)$$

where $W_{\mathrm{Pool}}^{t}(\cdot)$ denotes the level of social welfare in period t when the government uses pooling taxation. That is, the government chooses a single tax treatment $\langle c^t, y^t \rangle$ for both types to maximize

$$(1-\phi)\left[\frac{1}{1-\sigma}(c^t)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y^t}{w_{\mathrm{L}}}\right)^{1+\gamma}\right] + \phi\left[\frac{1}{1-\sigma}(c^t)^{1-\sigma} - \frac{1}{1+\gamma}\left(\frac{y^t}{w_{\mathrm{H}}}\right)^{1+\gamma}\right], \quad (21)$$

subject to the budget constraint

$$y^t - c^t \geq 0. \quad (22)$$

Because the budget constraint will be binding, the solution to program (21)–(22) will involve $c^t = y^t = y^t(\phi, \sigma, \gamma, w_{\mathrm{L}}, w_{\mathrm{H}})$. Substituting this function into (21) yields the level of social welfare in period t under pooling, that is, $W_{\mathrm{Pool}}^{t}(\cdot)$.

4. QUANTITATIVE ANALYSIS

As discussed earlier, it is not possible to derive our main results analytically. Accordingly, in this section we use numerical simulations to examine the effects of loose commitment. Following the standard practice for conducting numerical simulations, we first calibrate a benchmark version of our model, using generally accepted and empirically plausible parameter values. The OECD (2010) reports that on the average across OECD countries, approximately one-fourth of all adults

Table 1. Baseline parameter values

φ	0.250	δ	0.952	w_L	1.000
σ	1.000	r	0.050	w_H	1.600
γ	2.000	$p = q_s$	0.950		
		q_F	0.250		

have attained tertiary-level education. We therefore assume that 25% of individuals are high-skilled workers; i.e., we set $\phi = 0.25$. Fang (2006) and Goldin and Katz (2007) estimate that the college wage premium, i.e., the average difference between the wages of university graduates and high-school graduates, is approximately 60%. We therefore normalize the low-skilled type's wage to unity ($w_L = 1$) and set the high-skilled type's wage at $w_H = 1.6$. The preference parameter σ is set to unity, so that the utility function is logarithmic in consumption, and we set $\gamma = 2$, as this implies a labor supply elasticity of 0.5, which is broadly consistent with empirical estimates.[12] We assume that each period is 1 year in length and that the annual discount rate is 5%, which is in line with common practice. Finally, we assume that $p = q_S = 0.95$, in order to maintain the spirit of loose commitment; i.e., the probability of commitment is high, but not certain. However, we set q_F equal to the (relatively high) value of 0.25, in order to capture the idea that if the government does happen to use first-best taxation in period t, it is relatively more likely to use first-best taxation again in period $t + 1$. The baseline parameter values are presented in Table 1.

4.1. Benchmark Numerical Results

Figure 1 compares the level of social welfare attainable when separating taxation is used in period 1 and the level of social welfare in autarky,[13] for various values of p,[14] while all other parameters are held at their baseline levels. It can be seen that taxation with separation in period 1 yields a higher level of social welfare than autarky only when $p > 0.8$. However, it is theoretically possible that pooling taxation may do better. Figure 2 shows the level of social welfare with separation in period 1 minus the level of social welfare with separation in period T (thus pooling occurs for $T - 1$ periods), for various values of p. Separating in period 1 is worse than separating in period $T \geq 2$ only when p falls to around 0.5. Therefore, we conclude that the government can improve upon the free-market solution when the probability of commitment is greater than 80%, and in doing so, it is optimal for the government to separate the individuals in period 1. Accordingly, the remainder of our analysis is based on optimal taxation with separation occurring in the first period.

Next, Table 2 summarizes the optimal tax and welfare outcomes under loose commitment (using the baseline parameter values in Table 1) versus those under full commitment. Under loose commitment, the optimal average tax rate faced by high-skilled individuals in period 1 is *negative*, despite the government's

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

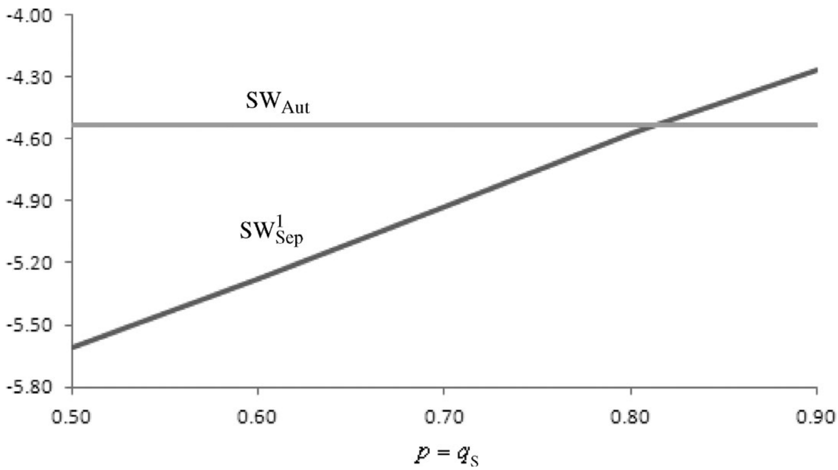

FIGURE 1. Social welfare with separation in period 1 versus social welfare under autarky.

redistributive concerns, whereas correspondingly that for low-skilled individuals is positive. The intuition is that high-skilled individuals know that revealing their type in period 1 will result in their facing first-best taxation in some periods in the future. Therefore, high-skilled individuals have to be compensated in the first period—which comes at the expense of low-skilled individuals—for the unfavorable tax treatments they will sometimes face after revealing their type.

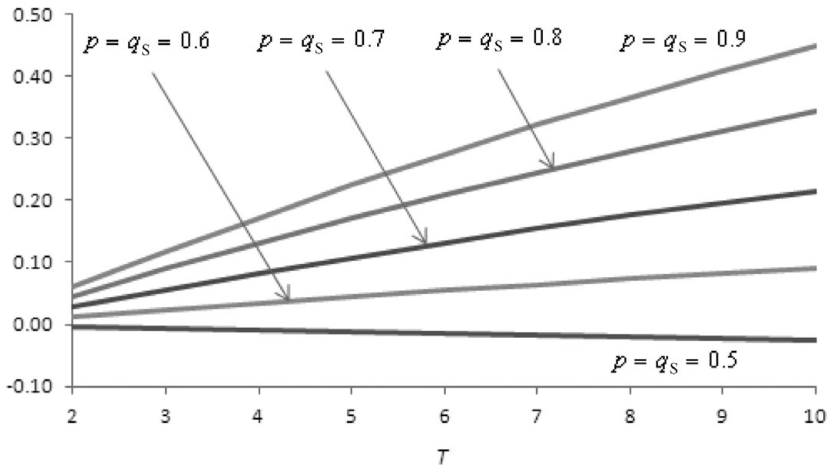

FIGURE 2. Social welfare with separation in period 1 minus social welfare with separation in period $T \geq 2$.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

TABLE 2. Loose commitment versus full commitment

	Loose commitment	Full commitment
First-period average tax rate: low-skilled	0.126	−0.091
First-period average tax rate: high-skilled	−0.249	0.153
First-period marginal tax rate: low-skilled	0.181	0.067
First-period marginal tax rate: high-skilled	0.000	0.000
First-period utility: low-skilled	−0.469	−0.250
First-period utility: high-skilled	0.351	−0.035
First-period social welfare	−0.264	−0.196
Lifetime utility: low-skilled	−5.301	−5.251
Lifetime utility: high-skilled	−0.776	−0.729
Lifetime social welfare	−4.170	−4.120

As in standard nonlinear income tax models, in period 1 high-skilled individuals face a zero marginal tax rate, whereas low-skilled individuals face a positive marginal tax rate. However, the low-skilled type's marginal tax rate is higher under loose commitment than under full commitment. This is because under loose commitment high-skilled individuals have a stronger incentive to mimic low-skilled individuals, in order to avoid facing first-best taxation. This makes it harder for the government to satisfy the high-skilled type's incentive-compatibility constraint. Accordingly, there is a greater need to distort the low-skilled type's labor supply downward through a positive marginal tax rate to relax the incentive-compatibility constraint.

In terms of the pattern of average tax rates, high-skilled individuals are better off in period 1 under loose commitment than under full commitment, whereas the opposite is true for low-skilled individuals. First-period social welfare is lower under loose commitment, because the government is forced to redistribute from low-skilled to high-skilled individuals. In the long run, however, both types of individual are better off and social welfare is higher under full commitment. This reflects the long-run benefits to be gained by a government that is able to commit.

Figure 3 shows the effects of variations in p around its baseline value on the first-period average tax rates, while all other parameters are held at their baseline levels. It can be seen that high-skilled individuals continue to face a negative average tax rate in period 1 unless it is almost certain that the government can commit, i.e., when $p > 98\%$. In this case, the compensation required by high-skilled individuals is not so severe that they need to be subsidized. Figure 4 shows the effects of varying p on the first-period marginal tax rates. As p increases, high-skilled individuals are more willing to reveal their type, which makes it easier for the government to satisfy their incentive-compatibility constraint. Accordingly, the low-skilled type's marginal tax rate can be reduced. Figure 5 shows the effects

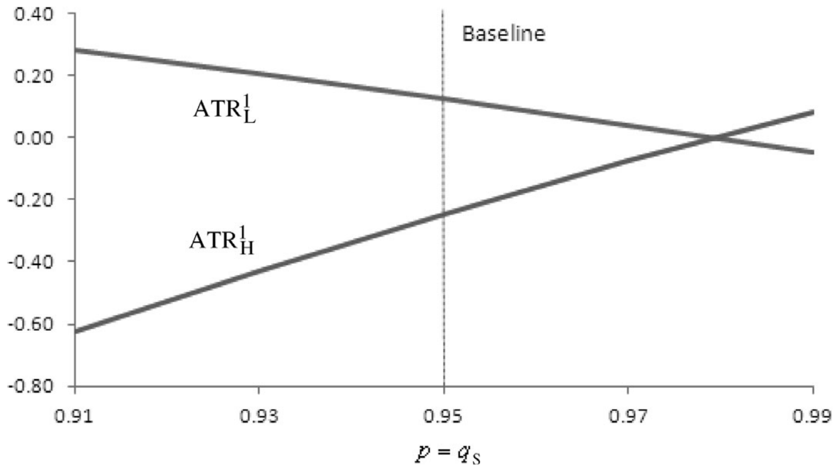

FIGURE 3. First-period average tax rates.

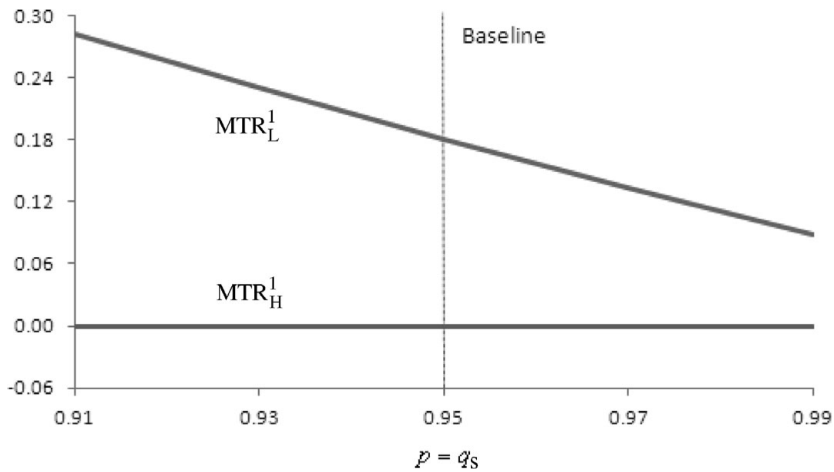

FIGURE 4. First-period marginal tax rates.

of varying p on first-period utility levels, which simply mirror the effects on first-period average tax rates. Figure 6 shows that the lifetime utility of both types of individuals is increasing in p, albeit only slightly, which reflects the long-run benefits of commitment.

4.2. Sensitivity Analysis

Our sensitivity analysis begins with Figure 7, illustrating the effects of varying the proportion of high-skilled individuals in the economy, ϕ, while the other

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

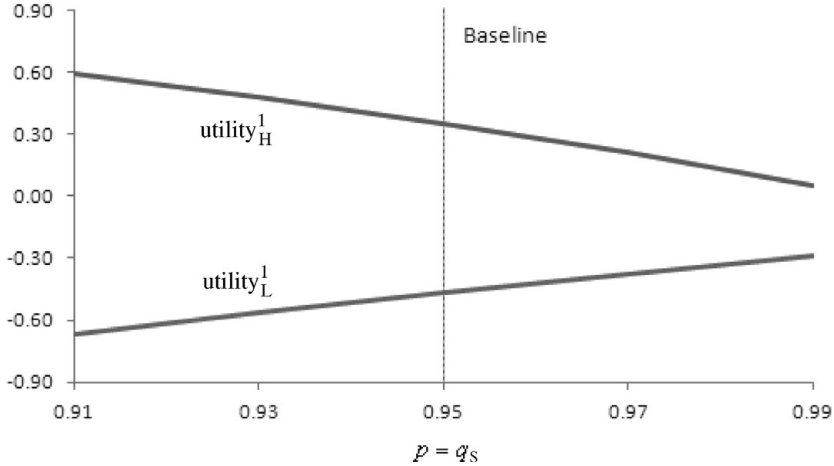

FIGURE 5. First-period utility.

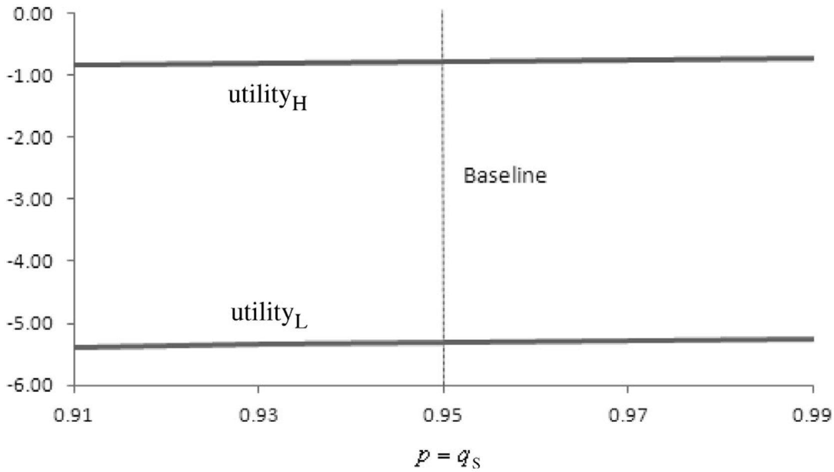

FIGURE 6. Lifetime utility.

parameters are held at their baseline levels. Simulations are conducted for the benchmark loose-commitment parameter values, as well as for the case of full commitment. The left column of Figure 7 shows the short-run (period 1) effects on individual welfare, whereas the right column shows the long-run (lifetime) effects. In the long run, both types of individual are better off as ϕ increases, whether or not commitment is certain, because the society is better off with a larger population of high-skilled individuals. This is because high-skilled individuals have a higher wage than low-skilled individuals, so an increase in the proportion of high-skilled

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

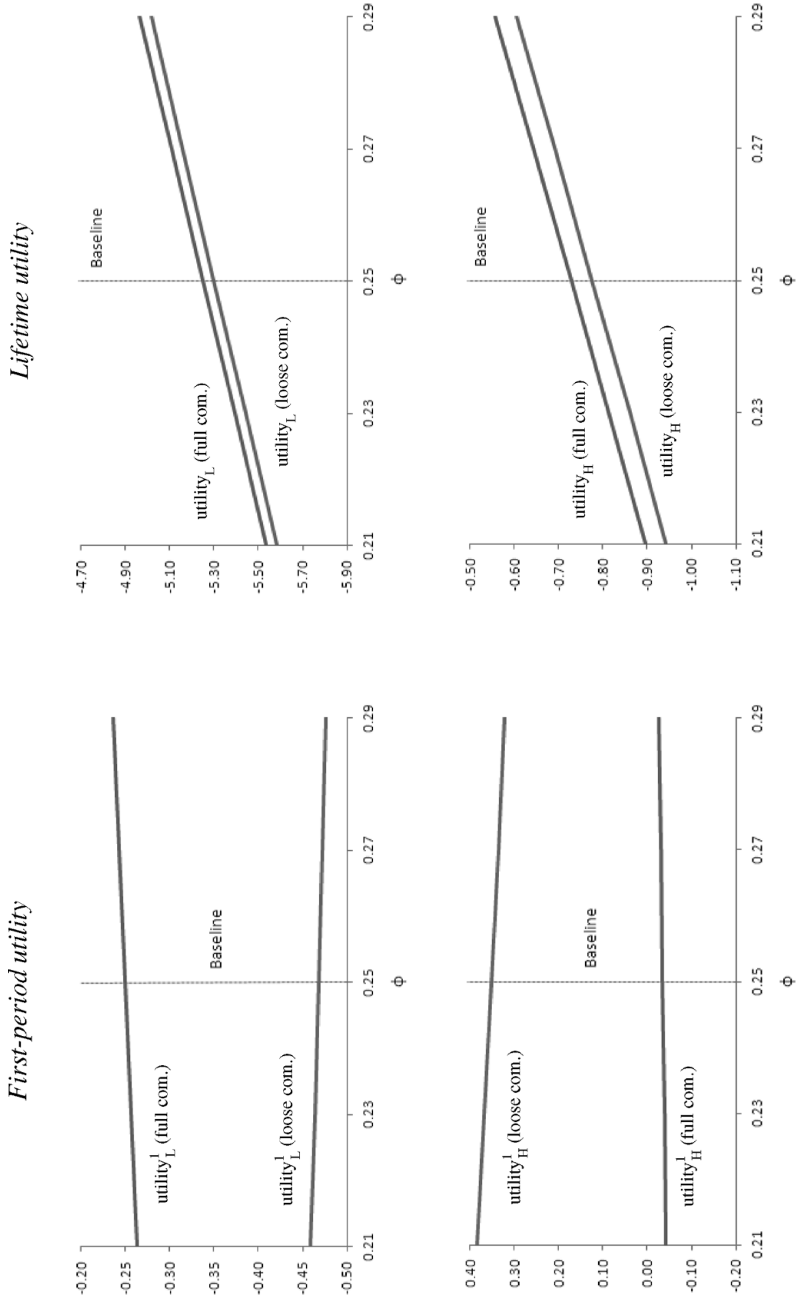

FIGURE 7. Effects of varying ϕ.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

individuals in the economy increases the economy's endowments, which enables the government to use its taxation powers to make everyone better off. Both types of individual are also better off in period 1 as ϕ rises under full commitment, but both types are *worse off* in period 1 as ϕ increases under loose commitment. Under first-best taxation, low-skilled individuals obtain a higher level of utility than high-skilled individuals. This is because under first-best taxation, both types receive the same level of consumption, but high-skilled individuals are required to work longer. However, an increase in ϕ reduces the difference in utility that low-skilled and high-skilled individuals obtain under first-best taxation, because the weight that high-skilled individuals receive in the social welfare function is higher. This makes mimicking less attractive, which in turn lowers the compensation—and hence utility—that high-skilled individuals require in period 1 to reveal their type. Low-skilled individuals, however, are also worse off in period 1 as ϕ increases. This is because under loose commitment, high-skilled individuals are subsidized in period 1 (as discussed earlier). An increase in their population therefore requires more taxation of each low-skilled individual to satisfy the government's first-period budget constraint; hence low-skilled individuals are also made worse off.

Figure 8 shows the effects of varying the high-skilled type's wage, w_H. In the long run, both types of individual are better off as w_H rises, under loose and full commitment, because a ceteris paribus increase in w_H corresponds to an increase in the economy's endowments. Both types are also better off in the short run as w_H increases, except for the low-skilled type under loose commitment. This is because individuals' utility is *decreasing* in their wage rate under first-best taxation, as they are required to work longer and do not receive a compensating increase in consumption. An increase in w_H therefore implies that high-skilled individuals require more compensation in period 1 to reveal their type, which comes at the expense of low-skilled individuals.

Figure 9 shows the effects of varying the discount rate, r. Simulations are again conducted for loose and full commitment, while all other parameters are held at their baseline levels. The lifetime utility of both types of individual is increasing in r, whether or not commitment is certain, simply because a lower discount factor is used to sum the infinite utility streams, and utility happens to be measured along the negative real line. Under full commitment, changes in r have no effect on either type's first-period utility, because the exact same allocation is implemented in each period and changes in r only affect the value of utility from period 2 onward. However, under loose commitment the low-skilled type's first-period utility is increasing in r, whereas that for the high-skilled type is decreasing. When r increases, high-skilled individuals discount the future at a greater rate, and therefore care less about the utility they obtain from period 2 onward. Accordingly, they require less compensation in period 1 to reveal their type.[15] This results in high-skilled individuals being worse off in period 1 as r increases, whereas low-skilled individuals are correspondingly made better off.

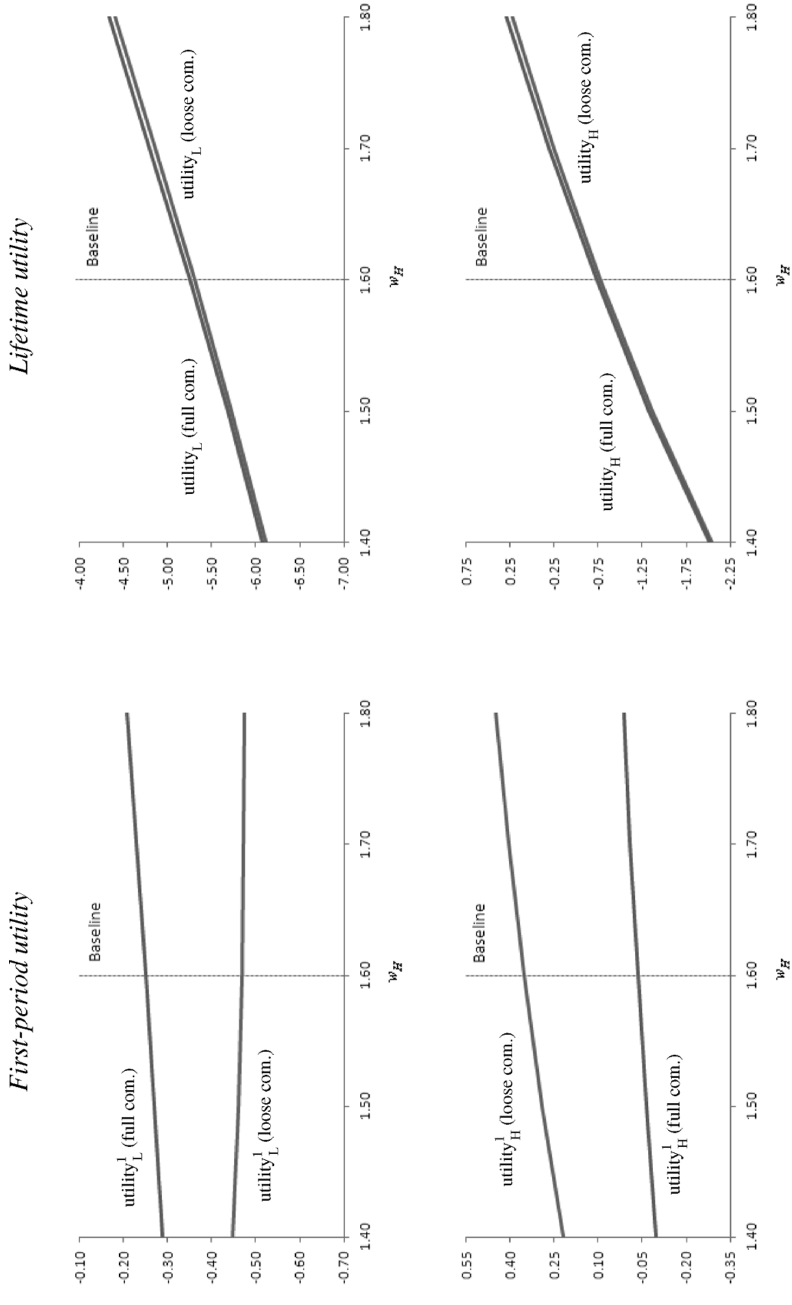

FIGURE 8. Effects of varying w_H.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

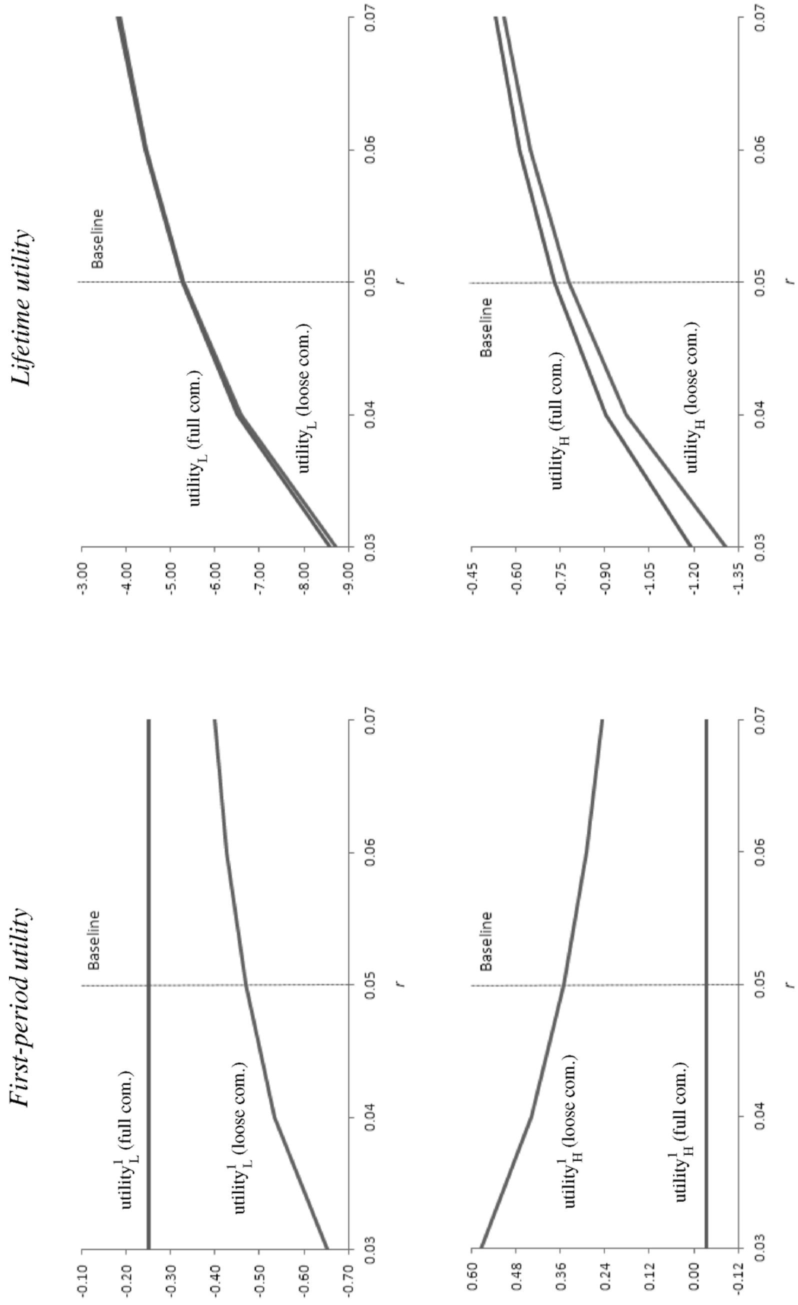

FIGURE 9. Effects of varying r.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

Figure 10 shows the effects of varying the labor supply elasticity, i.e., $1/\gamma$. We consider large changes in the labor supply elasticity, as micro-econometric estimates tend to yield low values, whereas macroeconomic estimates are significantly higher. In the long run, under full and loose commitment, both types of individual are worse off as the labor supply elasticity rises, because for the parameters of our model this corresponds to an increase in the disutility of labor. The same relationship holds true in the short run for the low-skilled type, but not for the high-skilled type. Specifically, the high-skilled type's first-period utility is *increasing* in $1/\gamma$ under loose commitment. The reason is that redistribution under first-best taxation becomes increasingly severe as the disutility of labor rises, which causes high-skilled individuals to demand more compensation to reveal their type. Therefore, their first-period utility is increasing in the disutility of labor, which provides another channel through which loose commitment can reverse the short-run welfare effects of parameter changes and yield counterintuitive outcomes.

Finally, Figures 11 and 12 show the effects of varying the Markov switching probabilities, q_S and q_F, while again all other parameters (including p) are held at their baseline levels. Under full commitment, changes in the Markov switching probabilities do not, of course, affect welfare. Under loose commitment, increases in q_S and decreases in q_F make both types of individual better off in the long run, which simply reflect the long-run benefits of moving toward full commitment. In the short run, increases in q_S make low-skilled individuals better off and high-skilled individuals worse off. As first-best taxation is now less likely to be implemented, high-skilled individuals require less compensation to reveal their type. Analogously, increases in q_F make low-skilled individuals worse off and high-skilled individuals better off in the short run, as the latter now demand more compensation to reveal their type.

5. CONCLUDING COMMENTS

Recent interest in dynamic nonlinear income taxation has raised the question of whether the government can commit to not take advantage of skill-type information revealed in earlier periods. This paper has assumed that there is only a very small probability that the government cannot commit. In this loose commitment setting, separating taxation remains optimal, as under full commitment. But nevertheless, loose commitment has a substantial impact on optimal dynamic nonlinear income taxation. Our quantitative analysis shows that even if commitment is almost certain, high-skilled individuals must be subsidized in the short run. We have also shown that loose commitment reverses almost all of the short-run welfare effects of changes in the model's parameters. The main message of our paper is, therefore, that even a little uncertainty over whether the government can commit has significant and counterintuitive effects.

Because this paper is, to the best of our knowledge, the first to analyze dynamic nonlinear income taxation in a loose commitment setting, we have studied the rather simplified two-type version of the Mirrlees model. In addition, we have

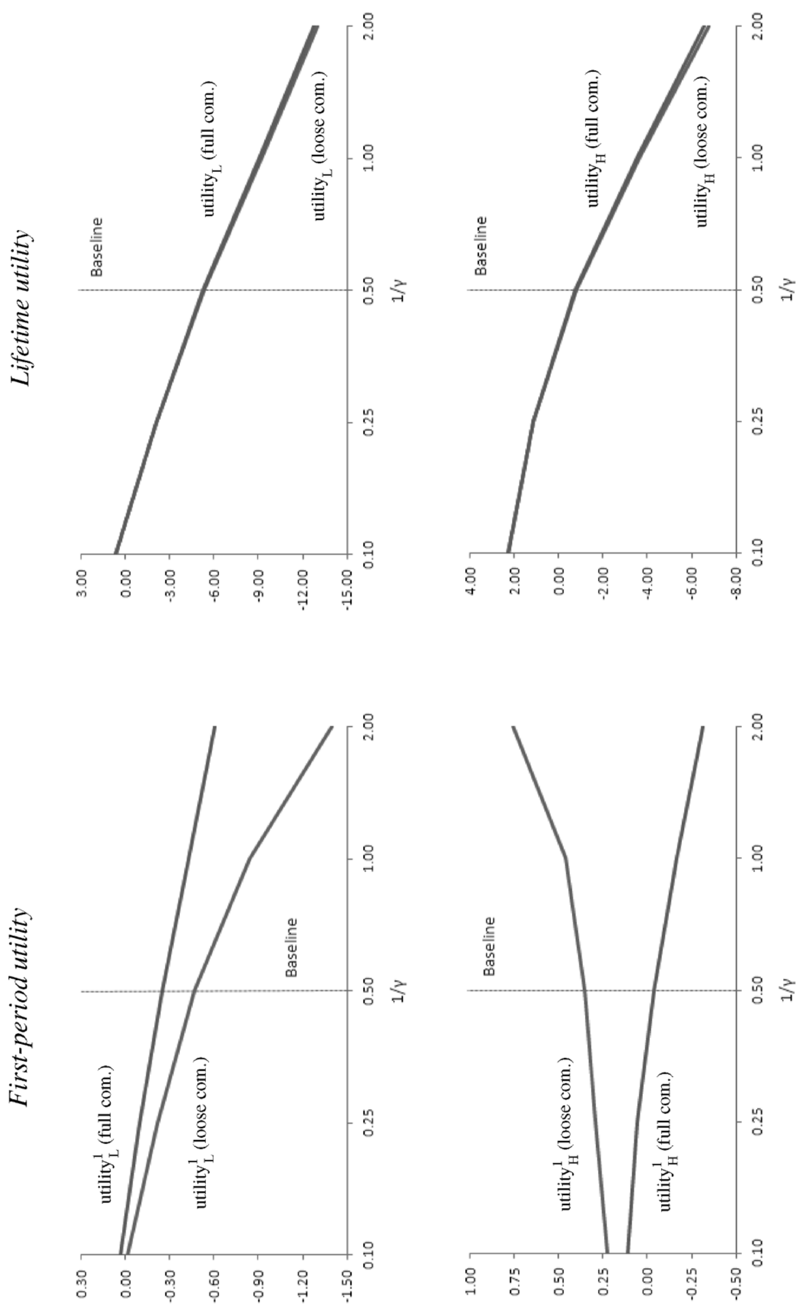

FIGURE 10. Effects of varying $1/\gamma$.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

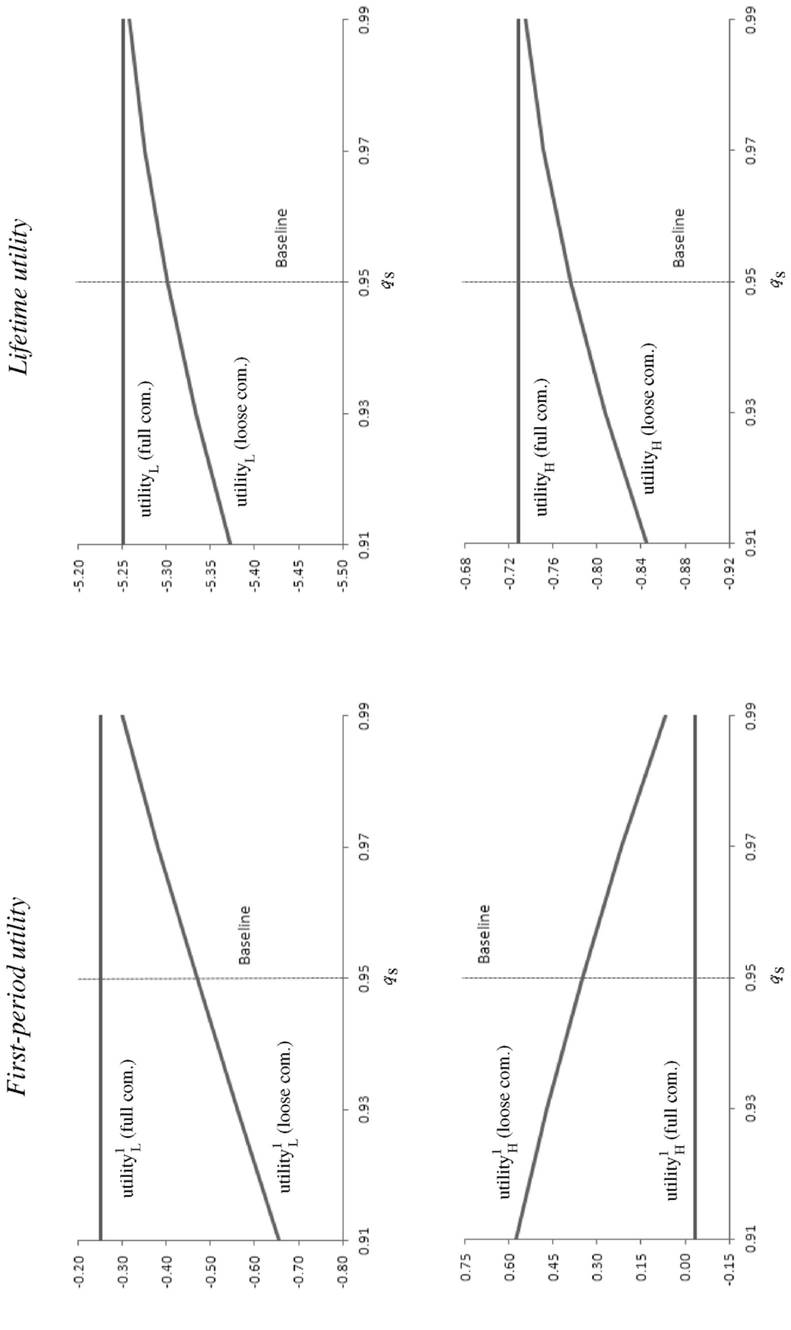

FIGURE 11. Effects of varying q_s.

https://doi.org/10.1017/S1365100512001010 Published online by Cambridge University Press

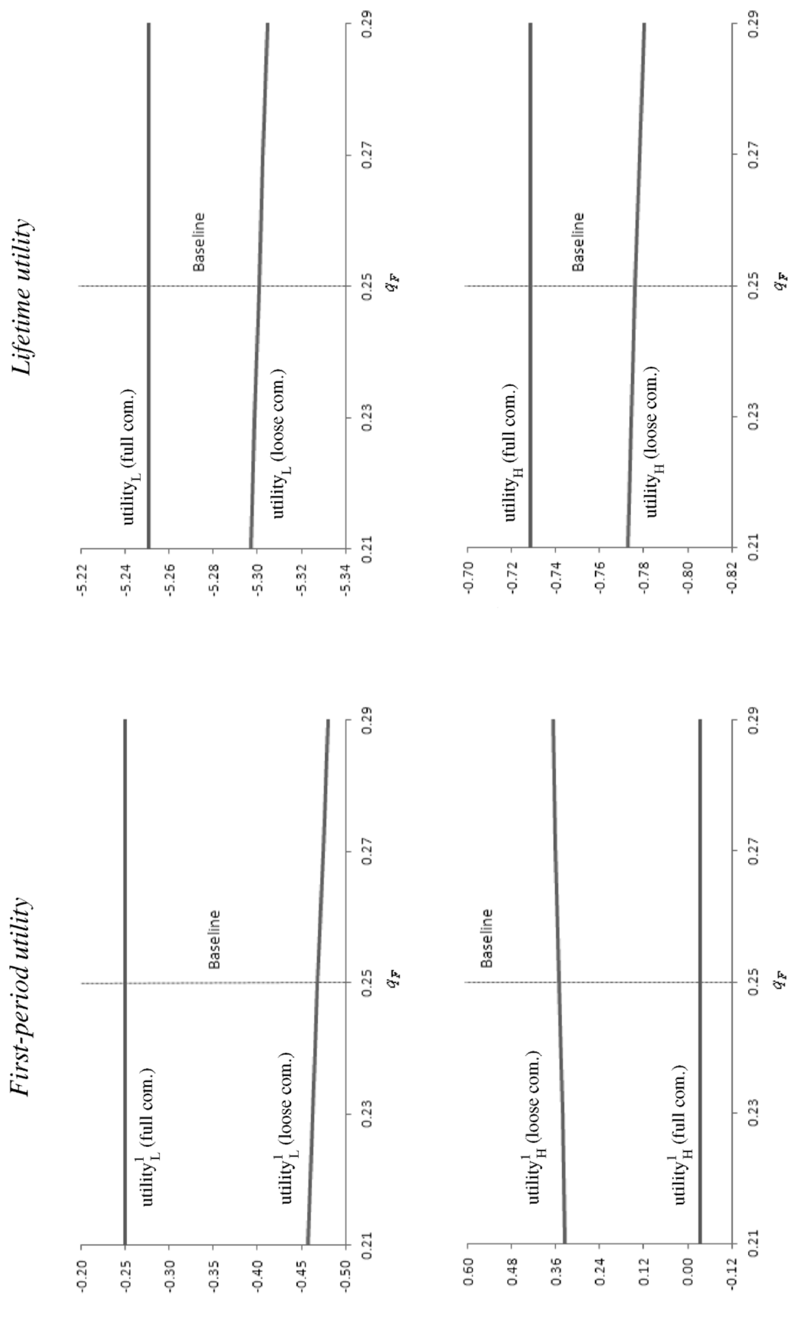

FIGURE 12. Effects of varying q_F.

assumed that the only link between periods is the revelation and possible use of skill-type information. However, as can be seen from our analysis, even extending this simple version of the Mirrlees model to an infinite-horizon setting with loose commitment leads to a fairly complicated optimal tax problem. It is also sufficient to bring out a number of interesting and counterintuitive results. That said, our model clearly has its limitations, and at least two potential extensions come to mind.

First, one could extend the model to a many-type setting, but we think our main conclusions would probably remain intact. What drives our results is that high-skilled individuals are better off under second-best taxation, whereas low-skilled individuals are better off under first-best taxation. If there are more than two types, then given the government's redistributive objective, there will still be one group of individuals, the higher-skilled, who are better off under second-best taxation, and another group, the lower-skilled, who are better off under first-best taxation. The main challenge for designing optimal dynamic nonlinear income taxation therefore remains essentially a two-group problem. Alternatively, one could stick with the two-type model, but allow individuals to change type, as in Battaglini and Coate (2008). However, as these authors show, the analysis then becomes extremely complex, and it would be even more so with loose commitment.

Second, our assumption that the only link between periods is the revelation and possible use of skill-type information allows us to isolate the effects of loose commitment from any other dynamic factors. An extension to a setting in which there are other dynamic links, such as public and private savings, is worth pursuing in future research. Nevertheless, this extension would make the analysis substantially more complicated, as an individual's utility would not depend only upon the government's history of use/nonuse of skill-type information; it would also depend upon the history of savings by individuals and the government. As in our model, high-skilled individuals would still feel the need for compensation in the short run if they are to reveal their type. But giving the government additional instruments, such as the ability to save and/or to tax private savings, may allow it to compensate high-skilled individuals in ways other than through short-run subsidization. However, it is difficult to conjecture as to how introducing savings would affect our results, because it would seem to depend upon the exact manner in which savings are introduced.

NOTES

1. Surveys of the new dynamic public finance literature are provided by Golosov et al. (2006, 2011). For a textbook treatment, see Kocherlakota (2010).

2. Gaube (2007) makes this argument.

3. The commitment assumption has also been criticized as being inconsistent with the microfoundations of the Mirrlees model. A key feature of the Mirrlees approach to optimal taxation is that no ad hoc constraints are placed on the tax instruments available to the government—these are determined only by the information structure. Thus ruling out lump-sum taxation in dynamic versions of the Mirrlees model via a commitment assumption might be considered inappropriate.

4. See, e.g., Roberts (1984), Berliant and Ledyard (2011), and Guo and Krause (2011a).

5. For a general theoretical treatment of dynamic contracting with imperfect commitment, see Bester and Strausz (2001). The term "loose commitment" is taken from Debortoli and Nunes (2010). They revisit the classic question of whether taxation should fall predominantly on capital or labor income within a prototypical dynamic representative-agent model, but one where the government can commit only with a certain probability.

6. See, e.g., Brett and Weymark (2008b, 2011) and Simula (2010).

7. Specifically, if the utility function is quasi-linear in labor (resp. consumption), then the first-best levels of pre-tax income (resp. consumption) cannot be uniquely determined.

8. There exists a third possibility in which the government plays a "mixed strategy" by pooling some, but not all, of the high-skilled individuals with the low-skilled individuals. As a first attempt at studying dynamic nonlinear income taxation under loose commitment, we restrict attention to the pure-strategy policies of complete separation or complete pooling taxation, and leave the case of partial pooling for future research.

9. Throughout the paper, we focus on the case often studied in the literature in which the government's revenue requirement is normalized to zero, so that the tax system is purely redistributive. If the government's revenue requirement were positive, then our main conclusion that the tax burden must fall predominantly on low-skilled individuals in the short run would remain intact, although high-skilled individuals would not necessarily be subsidized. Likewise, all of our results regarding the effects of changes in the model's parameters would be qualitatively the same.

10. The low-skilled type's incentive-compatibility constraint is not considered because the government will use its taxation powers to redistribute from high-skilled to low-skilled individuals under our model parameterizations. This creates an incentive for high-skilled individuals to "mimic" low-skilled individuals, but not vice versa. Accordingly, the high-skilled type's incentive-compatibility constraint will bind at an optimum, whereas the low-skilled type's incentive-compatibility constraint will be slack.

11. We continue to omit the low-skilled type's incentive-compatibility constraint because it will not be binding at an optimum.

12. See, e.g., Kocherlakota (2010, p. 189) and the review article by Chetty et al. (2011).

13. In autarky, each individual i will choose c_i^t and l_i^t to maximize equation (1) subject to the budget constraint $c_i^t \leq w_i l_i^t$ in each period.

14. As our baseline assumption is that $p = q_S$, changes in p also involve the same changes in q_S, but for brevity we simply refer to changes in p. In Subsection 4.2 we consider the effects of changing q_S independent of p.

15. For a similar reason, Berliant and Ledyard (2011) conclude that separating taxation remains optimal in a two-period Mirrlees model without commitment, provided the discount rate is sufficiently high.

REFERENCES

Apps, Patricia and Ray Rees (2006) Repeated Optimal Nonlinear Income Taxation. Mimeo, University of Munich.

Battaglini, Marco and Stephen Coate (2008) Pareto efficient income taxation with stochastic abilities. *Journal of Public Economics* 92, 844–868.

Berliant, Marcus and John Ledyard (2011) Optimal Dynamic Nonlinear Income Taxes with No Commitment. Mimeo, University of Washington, St. Louis.

Bester, Helmut and Roland Strausz (2001) Contracting with imperfect commitment and the revelation principle: The single agent case. *Econometrica* 69, 1077–1098.

Brett, Craig and John Weymark (2008a) Optimal Nonlinear Taxation of Income and Savings without Commitment. Working paper 08-W05, Vanderbilt University.

Brett, Craig and John Weymark (2008b) The impact of changing skill levels on optimal nonlinear income taxes. *Journal of Public Economics* 92, 1765–1771.

Brett, Craig and John Weymark (2011) How optimal nonlinear income taxes change when the distribution of the population changes. *Journal of Public Economics* 95, 1239–1247.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber (2011) Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review* 101, 471–475.

Debortoli, Davide and Ricardo Nunes (2010) Fiscal policy under loose commitment. *Journal of Economic Theory* 145, 1005–1032.

Fang, Hanming (2006) Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review* 47, 1151–1185.

Farhi, Emmanuel, Christopher Sleet, Ivan Werning, and Sevin Yeltekin (2012) Nonlinear capital taxation without commitment. *Review of Economic Studies* 79, 1469–1493.

Gaube, Thomas (2007) Optimum taxation of each year's income. *Journal of Public Economic Theory* 9, 127–150.

Goldin, Claudia and Lawrence Katz (2007) The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials 1890–2005. NBER working paper 12984.

Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski (2011) Optimal taxation: Merging micro and macro approaches. *Journal of Money*, *Credit and Banking* 43, 147–174.

Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning (2006) New dynamic public finance: A user's guide. *NBER Macroeconomics Annual* 2006, 317–387.

Guo, Jang-Ting and Alan Krause (2011a) Dynamic Income Taxation without Commitment: Comparing Alternative Tax Systems. Mimeo, University of California, Riverside.

Guo, Jang-Ting and Alan Krause (2011b) Optimal nonlinear income taxation with habit formation. *Journal of Public Economic Theory* 13, 463–480.

Guo, Jang-Ting and Alan Krause (2013) Optimal nonlinear taxation of income and education expenditures. *Oxford Economic Papers* 65, 74–95.

Kocherlakota, Narayana (2010) *The New Dynamic Public Finance*. Princeton, NJ: Princeton University Press.

Krause, Alan (2009) Optimal nonlinear income taxation with learning-by-doing. *Journal of Public Economics* 93, 1098–1110.

Mirrlees, James (1971) An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38, 175–208.

OECD (2010) *Highlights from Education at a Glance 2010*. Paris: OECD.

Roberts, Kevin (1984) The theoretical limits to redistribution. *Review of Economic Studies* 51, 177–195.

Simula, Laurent (2010) Optimal nonlinear income tax and nonlinear pricing: Optimality conditions and comparative static properties. *Social Choice and Welfare* 35, 199–220.

Stiglitz, Joseph (1982) Self-selection and Pareto efficient taxation. *Journal of Public Economics* 17, 213–240.