THE UNIVERSITY *of York*

Discussion Papers in Economics

Optimal Dynamic Nonlinear Income Taxation under Loose Commitment

By

Jang-Ting Guo, University of California, Riverside
Alan Krause, University of York

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

Optimal Dynamic Nonlinear Income Taxation under Loose Commitment[*]

Jang-Ting Guo[†]
University of California, Riverside

Alan Krause[‡]
University of York

October 14, 2010

Abstract

This paper examines an infinite-horizon model of dynamic nonlinear income taxation in which there exists a small probability that the government cannot commit to its future tax policy. 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. Under an empirically plausible parameterization, numerical simulations show that high-skill 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 model parameters. In particular, all individuals are worse-off, rather than better-off, in the short run when the proportion of high-skill individuals in the economy increases. Finally, our main findings remain qualitatively robust to a setting in which loose commitment is modelled as a Markov switching process.

Keywords: Dynamic Income Taxation; Loose Commitment.

JEL Classifications: H21; H24.

1 Introduction

There is currently a great deal of interest in dynamic nonlinear income taxation, as exemplified by the "new dynamic public finance" literature which 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 their skill type. In dynamic 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 onwards. In order to avoid this outcome, the new dynamic public finance literature typically assumes that the government can commit to its future tax policy.[2] 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,[3] and because 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, since the present government cannot easily impose binding constraints on the policies of future governments.[4] Accordingly, models of dynamic nonlinear income taxation without commitment have been developed by Apps and Rees [2006], Brett

---

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

[2] For example, the papers by Kocherlakota [2005], Albanesi and Sleet [2006] and Werning [2007] all assume that the government can commit.

[3] Gaube [2007] makes this argument.

[4] Auerbach [2006] cites the example of a proposal made to resolve the U.S. Social Security system's imbalance, which includes a tax increase to be made by the government in 2045. As Auerbach notes, such a proposal cannot be taken too seriously.

and Weymark [2008], Krause [2009], and Guo and Krause [2010a], among others. The structure of these models is relatively simple,[5] which allows them to make direct comparisons of optimal nonlinear income taxation with and without commitment. However, due mainly to the use of general preference formulations, attention has been restricted to comparing the signs of the optimal marginal tax rates,[6] which does not really expose the full extent to which optimal nonlinear income tax systems may differ when the government can and cannot commit.

Since the assumptions of commitment or no-commitment can be viewed as polar cases, in this paper we depart from the existing literature by assuming that the government can commit to its future tax policy with probability p, and therefore cannot commit with probability $(1 - p)$. However, when the government cannot commit, it is well-known that the "revelation principle" may no longer hold. That is, it may no longer be social-welfare maximizing for the government to implement (separating) nonlinear income taxation in which all individuals are willing to reveal their skill types.[7] Instead, it may be optimal to pool the individuals, by imposing the same tax treatment on everyone, so that skill-type information is not revealed. In order to avoid this (arguably uninteresting) possibility, we assume that p is sufficiently high so that separation in period 1 remains optimal; hence the term "loose commitment".[8] To the best of our knowledge, this paper is the *first* to examine dynamic nonlinear income taxation in a loose commitment framework. We use the two-type version of the Mirrlees model introduced by Stiglitz [1982], but extend it to an infinite-horizon setting. We further assume that preferences are quasi-linear in consumption. These simplifications lead to a model that lends itself readily to numerical simulations, which in turn allows us to investigate

---

[5]In particular, they all assume that there are only two types of individual, and that the economy has a finite (two-period) time horizon.

[6]Likewise, the new dynamic public finance literature has focused mainly on determining optimal "tax wedges", i.e., optimal marginal distortions.

[7]See, e.g., Roberts [1984], Berliant and Ledyard [2005], and Guo and Krause [2010b].

[8]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 where the government can commit only with a certain probability. Interestingly, these authors show that optimal linear capital and labor income taxes under loose commitment do not necessarily fall between the optimal levels that prevail in the polar cases of commitment and no-commitment.

in detail how loose commitment affects the characteristics of optimal dynamic nonlinear income taxation.

Our main finding is that even a small amount of uncertainty regarding whether the government can commit has a substantial effect on the design of an optimal dynamic nonlinear income tax system. In particular, under an empirically plausible parameterization, our quantitative results show that high-skill 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-skill individuals know that if the government cannot commit, they will forever face lump-sum taxation after revealing their type. Therefore, high-skill individuals require substantial compensation in period 1 if they are to reveal their type, even if the probability that the government cannot commit is negligible. Loose commitment also reverses the short-run welfare effects of changes in most parameters. For example, all individuals are worse-off, rather than better-off, in period 1 when the proportion of high-skill individuals in the economy increases. High-skill 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-skill individuals are also worse-off in period 1, because each low-skill individual must pay more tax to finance the larger total subsidy received by the increased population of high-skill individuals. The short-run welfare effects of varying wages, the weight low-skill individuals receive in the social welfare function, and the discount rate, are also shown to be affected—and often reversed—by loose commitment.

By simply assuming that the government can commit with probability p and cannot commit with probability $(1-p)$, our model has the desirable feature that it reduces to a model of dynamic nonlinear income taxation with commitment when $p = 1$, and to a model of dynamic nonlinear income taxation without commitment when $p = 0$. An alternative approach, however, is to assume that *in each period* there is some probability that the government can and cannot commit. Accordingly, we examine an extension of our model in which loose commitment is treated as a Markov switching process. That is, the probability that the government can commit (resp. not commit) in period $t + 1$ depends upon whether it implemented the commitment (resp. no-commitment)

tax system in period t (where $t \geq 2$). It is shown that all of our main conclusions obtained when the government can simply commit with probability p and not commit with probability $(1 - p)$ carry-over to this alternative setting.

The remainder of the paper is organized as follows. Section 2 describes the analytical framework that we consider, while 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. An extension of our model to a setting in which loose commitment is treated as a Markov switching process is examined in Section 5. Concluding comments are in Section 6, while two appendices contain some additional mathematical details.

2 Analytical Framework

There is a unit measure of infinitely-lived individuals, with a proportion $\phi \in (0, 1)$ being high-skill workers and the remaining $(1 - \phi)$ being low-skill workers. The high-skill type's wage is denoted by w_H, while that for the low-skill type is denoted by w_L, where $w_H > w_L$. For simplicity, wages are assumed to remain constant through time. The preferences of both types of individual in each period are represented by the (analytically convenient) quasi-linear in consumption utility function:

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

where c_i^t denotes type i's consumption in period t, l_i^t denotes type i's labor supply in period t, while $\alpha > 0$ and $\gamma > 0$ are preference parameters. 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 $y_i^t = w_i l_i^t$. We assume that individuals cannot save or borrow, which implies that $y_i^t - c_i^t$ is equivalent 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 measurable by a utilitarian social welfare function weighted towards low-skill

individuals. The government therefore has a redistributive objective, and it will be seeking to tax high-skill individuals in order to subsidize low-skill 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. Thus, in period 1, the government implements (the second-best) incentive-compatible nonlinear income taxation under which each individual is willing to reveal their skill type. The key question then is whether the government will use skill-type information revealed in period 1 to implement personalized lump-sum taxation from period 2 onwards. It is assumed that all individuals know that the government can commit with probability p, and therefore cannot commit with probability $(1 - p)$. That is, with probability p the government will continue to implement incentive-compatible taxation from period 2 onwards, but with probability $(1-p)$ the government will begin to implement personalized lump-sum taxation. As discussed earlier, it is assumed that p is sufficiently high so that separating taxation remains optimal. As with the individuals, for simplicity we assume that the government cannot save or borrow. Therefore, the only link between periods is the revelation and possible use of skill-type information by the government.

3   Optimal Taxation under Loose Commitment

At the end of period 1, the government has enough information to begin implementing first-best taxation. However, chance decides, with probabilities p and $(1 - p)$, whether the government can commit and therefore continues to implement second-best taxation, or whether the government cannot commit and therefore begins to implement first-best taxation. In this section, we start by describing optimal taxation from period 2 onwards when the commitment or no-commitment possibilities are realized. We then describe optimal taxation in period 1.

3.1   Second-Best Taxation from Period 2 onwards

If the government can commit, its behavior in period t ($t = 2, ..., \infty$) can be described as follows. Choose tax treatments $\langle c_L^t, y_L^t \rangle$ and $\langle c_H^t, y_H^t \rangle$ for the low-skill and high-skill

individuals, respectively, to maximize:[9]

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

subject to:

$$(1-\phi)(y_L^t - c_L^t) + \phi(y_H^t - c_H^t) \geq 0 \qquad (3.2)$$

$$\alpha c_H^t - \frac{1}{1+\gamma}\left(\frac{y_H^t}{w_H}\right)^{1+\gamma} \geq \alpha c_L^t - \frac{1}{1+\gamma}\left(\frac{y_L^t}{w_H}\right)^{1+\gamma} \qquad (3.3)$$

where equation (3.1) is the weighted utilitarian social welfare function, with $\pi \in (\frac{1}{2}, 1)$
being the weight the government attaches to the welfare of low-skill individuals.[10] Equa-
tion (3.2) is the government's budget constraint, which requires that total tax revenues be
non-negative. Equation (3.3) is the high-skill type's incentive-compatibility constraint.[11]
Even though the government knows, based on their first-period choices, each individual's
skill type, and therefore has enough information to implement personalized lump-sum
taxation, commitment implies that the government does not use this information. It
therefore continues to implement incentive-compatible taxation.

The solution to program $(3.1)-(3.3)$ yields the functions $c_L^t(\pi, \phi, \alpha, \gamma, w_L, w_H)$, $y_L^t(\cdot)$,
$c_H^t(\cdot)$ and $y_H^t(\cdot)$.[12] Substituting these functions into the utility function (2.1) yields $u_{iS}^t(\cdot)$,
which denotes the utility a type i individual obtains under incentive-compatible (i.e.,
second-best) taxation in period t $(t = 2, ..., \infty)$.

3.2   First-Best Taxation from Period 2 onwards

If the government cannot commit, it will take advantage of skill-type information re-
vealed in period 1 to implement personalized lump-sum taxation in periods $2, ..., \infty$.

---

[9]While we do not observe such tax treatments in practice, the revelation principle implies that
one can model the government as choosing each type's allocation directly subject to the budget and
incentive-compatibility constraints, rather than as specifying a nonlinear income tax system.

[10]Since we assume that the utility function is quasi-linear in consumption, we must impose the
restriction that $\pi > 0.5$ to ensure that the high-skill type's incentive-compatibility constraint is binding.

[11]The low-skill type's incentive-compatibility constraint is not considered because the government
will use its taxation powers to redistribute from high-skill to low-skill individuals under our model
parameterizations. This creates an incentive for high-skill individuals to "mimic" low-skill individuals,
but not vice versa. Accordingly, 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.

[12]Further details are provided in Appendix A.

In this case, the government can be described as choosing tax treatments $\langle \widetilde{c}_L^t, \widetilde{y}_L^t \rangle$ and $\langle \widetilde{c}_H^t, \widetilde{y}_H^t \rangle$ for the low-skill and high-skill individuals, respectively, to maximize:

$$\pi(1-\phi)\left[\alpha\widetilde{c}_L^t - \frac{1}{1+\gamma}\left(\frac{\widetilde{y}_L^t}{w_L}\right)^{1+\gamma}\right] + (1-\pi)\phi\left[\alpha\widetilde{c}_H^t - \frac{1}{1+\gamma}\left(\frac{\widetilde{y}_H^t}{w_H}\right)^{1+\gamma}\right] \qquad (3.4)$$

subject to:

$$(1-\phi)(\widetilde{y}_L^t - \widetilde{c}_L^t) + \phi(\widetilde{y}_H^t - \widetilde{c}_H^t) \geq 0 \qquad (3.5)$$

$$\widetilde{c}_H^t = \frac{1-\pi}{\pi}\widetilde{c}_L^t \qquad (3.6)$$

where the tilde decoration is used to distinguish first-best variables from their second-best counterparts. Equation (3.4) is the weighted utilitarian social welfare function, equation (3.5) is the government's budget constraint, and equation (3.6) is a constraint that is used to determine the first-best consumption levels. While quasi-linear preferences are, in general, analytically convenient, first-best taxation with such preferences is not completely determinate.[13] However, when preferences take the more general additively-separable form, $u(c_i^t) - v(l_i^t)$, with $u(\cdot)$ increasing and strictly concave and $v(\cdot)$ increasing and strictly convex, it is well known that first-best taxation under a strict utilitarian (i.e., $\pi = 0.5$) objective gives both types the same level of consumption. We therefore use this insight to determine the first-best consumption levels by adding constraint (3.6).[14]

The solution to program $(3.4)-(3.6)$ yields the functions $\widetilde{c}_L^t(\pi, \phi, \alpha, \gamma, w_L, w_H)$, $\widetilde{y}_L^t(\cdot)$, $\widetilde{c}_H^t(\cdot)$ and $\widetilde{y}_H^t(\cdot)$.[15] Substituting these functions into the utility function (2.1) yields $u_{iF}^t(\cdot)$, which denotes the utility a type i individual obtains under first-best taxation in

---

[13] This has been noted previously by others, e.g., Brett and Weymark [2008]. Specifically, when the utility function is quasi-linear in consumption (resp. labor), the first-best levels of consumption (resp. pre-tax income) cannot be determined.

[14] While the addition of constraint (3.6) to make first-best taxation with quasi-linear in consumption preferences fully determinate might be considered a little ad hoc, it will be shown later that (3.6) has no real bearing on our main conclusions. This is because our formulation preserves all the salient features of first-best taxation generally, namely: (i) both types receive the same level of consumption if $\pi = 0.5$, and low-skill individuals receive more consumption if $\pi > 0.5$, (ii) high-skill individuals work longer than low-skill individuals, (iii) high-skill individuals obtain less utility than low-skill individuals, and (iv) an individual's utility is decreasing in their wage rate.

[15] Further details are provided in Appendix A.

period t $(t = 2, ..., \infty)$.

3.3   Optimal Taxation in Period 1

In period 1, the government cannot distinguish high-skill from low-skill individuals. Moreover, each individual knows that after they reveal their type in period 1, the government will solve program $(3.1) - (3.3)$ in periods $2, ..., \infty$ with probability p, and will solve program $(3.4) - (3.6)$ in periods $2, ..., \infty$ with probability $(1-p)$. The government's behavior in period 1 can therefore be described as follows. Choose 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 maximize:

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

subject to:

$$(1-\phi)(y_L^1 - c_L^1) + \phi(y_H^1 - c_H^1) \geq 0 \qquad (3.8)$$

$$\alpha c_H^1 - \frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} + \sum_{t=2}^{\infty}\delta^{t-1}\left[pu_{HS}^t(\cdot) + (1-p)u_{HF}^t(\cdot)\right] \geq$$

$$\alpha c_L^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} + \sum_{t=2}^{\infty}\delta^{t-1}\left[pu_{HS}^t(\cdot) + (1-p)u_{LF}^t(\cdot)\right] \qquad (3.9)$$

where equation (3.7) is the first-period weighted utilitarian social welfare function, equation (3.8) is the government's first-period budget constraint, and equation (3.9) is the high-skill type's incentive-compatibility constraint.[16] In order for a high-skill individual to be 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 utility a high-skill individual can expect to obtain from period 2 onwards, cannot be less than the utility a high-skill individual could obtain by pretending to be low skill by choosing $\langle c_L^1, y_L^1 \rangle$ in period 1, plus the utility this "mimicking" high-skill individual can expect to obtain from period 2 onwards. Specifically, a high-skill individual who reveals their type in period 1 can expect to obtain $u_{HS}^t(\cdot)$ in periods $2, ..., \infty$ with probability p, and $u_{HF}^t(\cdot)$ in periods $2, ..., \infty$ with probability $(1-p)$. A mimicking high-skill individual can also expect to obtain $u_{HS}^t(\cdot)$

---

[16]We again omit the low-skill type's incentive-compatibility constraint, because it will not be binding at an optimum.

in periods $2, ..., \infty$ with probability p, because if the government can commit, a mimicking high-skill individual might as well reveal their true type from period 2 onwards as there is no advantage to be had from continuing to pretend to be low skill.[17] But if the government cannot commit, a mimicking high-skill individual will continue to be treated as low skill by the government, and in any event they are better-off under this arrangement since $u_{LF}^t(\cdot) > u_{HF}^t(\cdot)$.[18] A mimicking high-skill individual can therefore expect to obtain $u_{LF}^t(\cdot)$ in periods $2, ..., \infty$ with probability $(1 - p)$.

The solution to program $(3.7) - (3.9)$ yields the functions $c_L^1(\pi, \phi, \alpha, \gamma, w_L, w_H, \delta, p)$, $y_L^1(\cdot)$, $c_H^1(\cdot)$ and $y_H^1(\cdot)$.[19] Substituting these functions into the utility function (2.1) yields $u_{iS}^1(\cdot)$, which denotes the utility a type i individual obtains under second-best taxation in period 1.

Finally, the lifetime utility of a type i individual is equal to:

$$u_{iS}^1(\cdot) + p \sum_{t=2}^{\infty} \delta^{t-1} u_{iS}^t(\cdot) + (1 - p) \sum_{t=2}^{\infty} \delta^{t-1} u_{iF}^t(\cdot) \tag{3.10}$$

That is, an individual's lifetime utility is equal to the utility they obtain in period 1 plus the utility they can expect to obtain from period 2 onwards.

4 Numerical Simulations

In order to investigate how loose commitment affects optimal dynamic nonlinear income taxation, this section undertakes numerical simulations. We first identify a set of baseline parameter values for our model that are empirically plausible. The OECD [2010] reports that on average across OECD countries, approximately one-quarter of all adults have

---

[17]It is possible for a high-skill individual to mimic in period 1, and then reveal their true type from period 2 onwards, because under commitment the government ignores all skill-type information revealed in period 1.

[18]The intuition for $u_{LF}^t(\cdot) > u_{HF}^t(\cdot)$ is closely related to the standard result that first-best taxation under a strict utilitarian objective has an individual's utility decreasing in their wage rate, since all individuals receive the same level of consumption but higher-wage individuals are required to work longer. Indeed, this has led some to describe first-best taxation as Marxist in nature, as it takes from each individual according to their ability and gives to each individual according to their need.

[19]Further details are provided in Appendix A.

attained tertiary level education. We therefore assume that 25% of individuals are high-skill 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 in the wages of university graduates over high-school graduates, is approximately 60%. We therefore normalize the low-skill type's wage to unity ($w_L = 1$) and set the high-skill type's wage at $w_H = 1.6$.

The remaining parameters are chosen on the following basis. If $\pi = 0.5$, the social welfare function is strictly utilitarian, which is a standard assumption in the literature. However, since we are assuming quasi-linear in consumption preferences, we must impose the restriction that $\pi > 0.5$ to ensure that the high-skill type's incentive-compatibility constraint is binding. We therefore set $\pi = 0.55$ so that our weighted utilitarian social welfare function is approximately strictly utilitarian. The parameters of the utility function, α and γ, are both simply set to unity so that the utility function (2.1) is quadratic in hours worked. We assume that each period is one-year in length and that the annual discount rate is 5%, which is in line with common practice. Finally, we assume that there is a high probability that the government can commit, $p = 0.95$, which ensures that separating taxation remains optimal.[20] The baseline parameter values are presented in Table 1.

Table 1 also shows that the baseline utility low-skill individuals obtain in period 1 is almost ten times lower than that obtained by high-skill individuals, whereas the lifetime utility differential is less than double. This is because high-skill individuals require substantial compensation in period 1 for the possibility that they may face first-best taxation from period 2 onwards. This possibility is investigated further in Figure 1, which shows the effects of varying the probability of commitment, p, whilst holding all other parameters at their baseline levels. The top panel of Figure 1 shows that the optimal average tax rate faced by high-skill individuals in period 1 is *negative*, despite the government's redistributive concerns, unless $p > 98\%$. The intuition is that high-skill individuals know that if the government cannot commit, revealing their type in

---

[20]Separating taxation is not feasible in our model for all $p < 0.87$, as it would require that low-skill individuals face an average tax rate of more than 100% in the first period (holding all other parameters at their baseline levels).

11

period 1 will result in them facing first-best taxation from period 2 onwards. Therefore, high-skill individuals have to be compensated in the first period—which comes at the expense of low-skill individuals—for the unfavorable tax treatments they will face after revealing their type. This result will continue to hold until it is almost certain that the government can commit when $p > 98\%$. In this case, the compensation required by high-skill individuals is not so severe that they need to be subsidized in the first period.

The middle panel of Figure 1 shows the effects of varying p on first-period utility levels. Reflecting the effects on first-period average tax rates, the high-skill type's first-period utility is decreasing in p, while that for the low-skill type is increasing. The bottom panel of Figure 1, however, shows that the lifetime utility of high-skill individuals is increasing in p, while that for low-skill individuals is decreasing. Since in the long run high-skill individuals are better-off under second-best taxation and low-skill individuals are better-off under first-best taxation, an increase in p raises the lifetime utility of high-skill individuals and decreases that of low-skill individuals.

Figure 2 illustrates the effects of varying the proportion of high-skill individuals in the economy, ϕ, whilst holding the other parameters at their baseline levels. Simulations are conducted for the baseline loose commitment probability of $p = 0.95$, as well as for the case when commitment is certain ($p = 1$). The left column of Figure 2 shows the short-run (period 1) effects on individual welfare, while 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 society is better-off with a larger population of high-skill individuals. That is, because high-skill individuals have a higher wage than low-skill individuals, an increase in the proportion of high-skill 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 increases if commitment is certain, but both types are *worse-off* in period 1 as ϕ increases under loose commitment. An increase in ϕ reduces the difference in utility that low-skill and high-skill individuals obtain under first-best taxation, because the effective weight that high-skill individuals receive in the social welfare function, namely $(1-\pi)\phi$, increases. This makes mimicking less attractive,

which in turn lowers the compensation—and hence utility—that high-skill individuals require in period 1 to reveal their type. Low-skill individuals, however, are also worse-off in period 1 as ϕ increases. This is because when $p = 0.95$, high-skill individuals are subsidized in period 1 (as discussed earlier). An increase in their population therefore requires more taxation of each low-skill individual in order to satisfy the government's first-period budget constraint; hence low-skill individuals are also made worse-off.

Figure 3 shows the effects of varying the wage premium, w_H/w_L. In the long run, both types of individual are better-off as the wage premium increases, whether or not commitment is certain, simply because an increase in w_H/w_L corresponds to an increase in the economy's endowments. Both types are also better-off in the short run as the wage premium increases, except for the low-skill type under loose commitment. This is because under first-best taxation, the high-skill type's utility is decreasing in the wage premium.[21] Accordingly, an increase in the wage premium raises the compensation that high-skill individuals require in the first period to reveal their type. This compensation comes at the expense of low-skill individuals, who are therefore worse-off in period 1.

Figure 4 shows the effects of varying the weight low-skill individuals receive in the social welfare function, π. In the long run, low-skill individuals are better-off, and high-skill individuals are worse-off, as π rises, whether or not commitment is certain, simply because the government cares relatively more about the welfare of low-skill individuals. This pattern of welfare effects is also obtained in the short run when commitment is certain, although the numerical simulations suggest that the effects are negligible. However, under loose commitment low-skill individuals are worse-off, and high-skill individuals are better-off, in the short run when π increases. The intuition is that redistribution under first-best taxation becomes especially severe as π increases, thus making high-skill individuals worse-off. This increases the compensation they require in period 1 to reveal their type, which comes at the expense of low-skill individuals. Accordingly, high-skill individuals are better-off, and low-skill individuals are worse-off, in the short run under loose commitment when π increases.

---

[21]The intuition for this result is analogous to that underlying the standard result whereby an individual's utility under first-best taxation is decreasing in their wage rate.

Figure 5 shows the effects of varying the discount rate, r. Simulations are again conducted for the baseline loose commitment probability ($p = 0.95$), and for the case when commitment is certain ($p = 1$), while holding all other parameters at their baseline levels. The lifetime utility of both types of individual is decreasing in r, whether or not commitment is certain, simply because a lower discount factor is used to sum the infinite utility streams. Similarly, when commitment is certain, 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 onwards. However, under loose commitment the low-skill type's first-period utility is increasing in r, while that for the high-skill type is decreasing. When r increases, high-skill individuals discount the future at a greater rate, and therefore care less about the utility they obtain from period 2 onwards. Accordingly, they require less compensation in period 1 to reveal their type. This results in high-skill individuals being worse-off in period 1 as r increases, while low-skill individuals are correspondingly made better-off.

5   Loose Commitment as a Markov Process

Thus far, we have simply assumed that the government can commit with probability p and cannot commit with probability $(1 - p)$. The advantage of this approach is that, when $p = 1$ and $p = 0$, our model collapses to the polar cases of dynamic nonlinear income taxation with and without commitment, respectively. Furthermore, variations in p capture changes in the expectation that either the commitment or no-commitment tax systems will be implemented. However, an alternative approach is to assume that after skill-type information is revealed in period 1, there is some chance *in each period* from period 2 onwards that the government can and cannot commit. Accordingly, in this section we model loose commitment as a Markov switching process.

Specifically, suppose as before that with probability p the government implements the second-best allocation in period 2, and with probability $(1 - p)$ the government implements the first-best allocation in period 2. But now, from period 3 onwards, the probability that the government will implement the second-best or first-best allocation

14

is determined according to the following transition probabilities:

$$\Pr(SB \text{ in period } t+1 | SB \text{ in period } t) = q_S \tag{5.1}$$

$$\Pr(FB \text{ in period } t+1 | FB \text{ in period } t) = q_F \tag{5.2}$$

That is, if the government implements the second-best (SB) allocation in period t (where $t \geq 2$), there is a probability of q_S that it will implement the second-best allocation again in period $t+1$, and a probability of $(1-q_S)$ that it will switch and implement the first-best (FB) allocation in period $t+1$. Likewise, if the government implements the first-best allocation in period t (where $t \geq 2$), it implements the first-best allocation again in period $t+1$ with probability q_F, and it implements the second-best allocation in period $t+1$ with probability $(1-q_F)$.

For each period $t \geq 2$, the "continuation utility" of a type i individual can be written as the following recursive equations:

$$V_{iS}^t = u_{iS}^t(\cdot) + q_S \delta V_{iS}^{t+1} + (1-q_S)\delta V_{iF}^{t+1} \tag{5.3}$$

$$V_{iF}^t = u_{iF}^t(\cdot) + q_F \delta V_{iF}^{t+1} + (1-q_F)\delta V_{iS}^{t+1} \tag{5.4}$$

where V_{iS}^t (resp. V_{iF}^t) is type i's continuation utility if the second-best (resp. first-best) allocation is implemented in period t (where $t \geq 2$). For example, equation (5.3) can be interpreted as follows. If the government implements the second-best allocation in period t, a type i individual obtains their second-best utility level $u_{iS}^t(\cdot)$ in period t, and with probability q_S they continue to obtain their second-best utility level in period $t+1$, but with probability $(1-q_S)$ the government switches and they obtain their first-best utility level in period $t+1$. That is, the continuation utility function V_{iF}^{t+1} becomes active. Similarly, the continuation utility of a mimicking high-skill individual can be written as:

$$V_{MS}^t = u_{HS}^t(\cdot) + q_S \delta V_{MS}^{t+1} + (1-q_S)\delta V_{MF}^{t+1} \tag{5.5}$$

$$V_{MF}^t = u_{LF}^t(\cdot) + q_F \delta V_{MF}^{t+1} + (1-q_F)\delta V_{MS}^{t+1} \tag{5.6}$$

for all periods $t \geq 2$.

Since $u_{iS}^t(\cdot)$ and $u_{iF}^t(\cdot)$ are time invariant for all $t \geq 2$, one may drop the time superscripts in equations $(5.3) - (5.6)$ and solve for the functions $V_{iS}(\delta, q_S, q_F, u_{iS}, u_{iF})$, $V_{iF}(\delta, q_S, q_F, u_{iS}, u_{iF})$, $V_{MS}(\delta, q_S, q_F, u_{HS}, u_{LF})$ and $V_{MF}(\delta, q_S, q_F, u_{HS}, u_{LF})$.[22]

The government's behavior in period 1 when loose commitment is modelled as a Markov process can now be described as follows. Choose 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 maximize:

$$\pi(1-\phi) \left[\alpha c_L^1 - \frac{1}{1+\gamma} \left(\frac{y_L^1}{w_L} \right)^{1+\gamma} \right] + (1-\pi)\phi \left[\alpha c_H^1 - \frac{1}{1+\gamma} \left(\frac{y_H^1}{w_H} \right)^{1+\gamma} \right] \quad (5.7)$$

subject to:

$$(1-\phi)(y_L^1 - c_L^1) + \phi(y_H^1 - c_H^1) \geq 0 \quad (5.8)$$

$$\alpha c_H^1 - \frac{1}{1+\gamma} \left(\frac{y_H^1}{w_H} \right)^{1+\gamma} + p\delta V_{HS}(\cdot) + (1-p)\delta V_{HF}(\cdot) \geq$$

$$\alpha c_L^1 - \frac{1}{1+\gamma} \left(\frac{y_L^1}{w_H} \right)^{1+\gamma} + p\delta V_{MS}(\cdot) + (1-p)\delta V_{MF}(\cdot) \quad (5.9)$$

where equation (5.7) is the first-period weighted utilitarian social welfare function, equation (5.8) is the government's first-period budget constraint, and equation (5.9) is the high-skill type's incentive-compatibility constraint. A high-skill individual who reveals their type in period 1 by choosing $\langle c_H^1, y_H^1 \rangle$ can expect to obtain a continuation utility from period 2 of $V_{HS}(\cdot)$ with probability p, and $V_{HF}(\cdot)$ with probability $(1-p)$. On the other hand, a high-skill individual who mimics in period 1 by choosing $\langle c_L^1, y_L^1 \rangle$ can expect to obtain a continuation utility from period 2 of $V_{MS}(\cdot)$ with probability p, and $V_{MF}(\cdot)$ with probability $(1-p)$. Therefore, in order to induce high-skill individuals to reveal their type in period 1, the tax treatments must satisfy equation (5.9).

The solution to program $(5.7)-(5.9)$ yields $c_L^1(\pi, \phi, \alpha, \gamma, w_L, w_H, \delta, p, V_{HS}, V_{HF}, V_{MS}, V_{MF})$, $y_L^1(\cdot)$, $c_H^1(\cdot)$ and $y_H^1(\cdot)$.[23] Substituting these functions into the utility function (2.1) yields $u_{iS}^1(\cdot)$, which denotes the utility a type i individual obtains under second-best taxation

---

[22]Further details are provided in Appendix B.
[23]Further details are provided in Appendix B.

in period 1 when loose commitment is modelled as a Markov process.

Finally, the lifetime utility of a type i individual when loose commitment is modelled as a Markov process is equal to:

$$u_{iS}^1(\cdot) + p\delta V_{iS}(\cdot) + (1-p)\delta V_{iF}(\cdot) \tag{5.10}$$

That is, an individual's lifetime utility is equal to the utility they obtain in period 1 plus the continuation utility they can expect from period 2.

Next, we use the baseline parameter values in Table 1 to conduct numerical simulations for the Markov model, except we must now also specify values for the transition probabilities q_S and q_F. We assume that $q_S = p = 0.95$, 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 implement the first-best allocation in period t, it is relatively more likely to implement the first-best allocation again in period $t+1$.[24]

As it turns out, the simulation results when loose commitment is modelled as a Markov process are qualitatively identical to those in Figures $1-5$ when there is simply a probability p that the government can commit and a probability $(1-p)$ that the government cannot commit. Indeed, it is still the case that high-skill individuals must be subsidized in period 1 unless the probability of commitment $(p = q_S)$ is greater than 98%. The welfare effects of varying the remaining parameters, for the cases when commitment is certain $(p = q_S = 1)$ and for the baseline loose commitment probability $(p = q_S = 0.95)$, are also qualitatively the same as before and therefore are not presented here.[25] In sum, all of our main findings obtained earlier remain robust to modelling loose commitment as a Markov switching process.

---

[24]The baseline transition probabilities of $q_S = 0.95$ and $q_F = 0.25$ are also chosen on the basis that they ensure that separating taxation remains feasible.

[25]Details of the numerical simulations when loose commitment is modelled as a Markov switching process are available upon request.

6 Concluding Comments

Recent interest in dynamic nonlinear income taxation has raised the question of whether the government can commit to its future tax policy. This paper has shown, perhaps surprisingly, that a little uncertainty over whether the government can commit has a substantial impact on the design of an optimal dynamic nonlinear income tax system. Our quantitative results show that even if commitment is almost certain, high-skill individuals must be subsidized in the short run, despite the government's redistributive objective, in order to compensate them for the possibility that they may forever face lump-sum taxation after revealing their type. We have also shown that loose commitment reverses almost all of the short-run welfare effects of changes in the model's parameters. Finally, we have obtained these conclusions when the government can simply commit with probability p and not commit with probability $(1 - p)$, or when loose commitment is modelled as a Markov switching process.

In order to undertake a detailed investigation into how loose commitment affects optimal dynamic nonlinear income taxation, we have used a simple model that lends itself readily to numerical simulations. The question remains as to how dependent our results are on the specifics of the model. However, based on the nature of the intuition driving our results, we think that most would continue to hold in more general settings, e.g., under different preference assumptions. This is because most of our arguments do not depend upon the specifics of the model, but instead rely on the general fact that low-skill individuals are better-off under first-best taxation and high-skill individuals are better-off under second-best taxation. Likewise, if the model were extended to more than two skill types, we think our main conclusions would remain intact because an individual's utility is increasing in wages under second-best taxation, but decreasing in wages under first-best taxation. Thus higher-wage types will prefer second-best taxation, and lower-wage types will prefer first-best taxation. Nevertheless, extending the model to settings with different preferences and more than two skill types present worthwhile avenues for future research.

Other possible extensions include making the probability of commitment endogenous.

For example, since low-skill individuals are better-off under first-best taxation and high-skill individuals are better-off under second-best taxation, each type has an incentive to petition the government to use either first-best or second-best taxation. Accordingly, extending our model to include some sort of lobbying behavior by individuals would be interesting, although it would change the analysis from being purely normative in nature to also being partly positive. Finally, the only link between periods in our model is the revelation and possible use of skill-type information by the government. An extension to a setting in which there are other dynamic links, such as savings, would also be worth pursuing.

7 Appendix A

Program $(3.1) - (3.3)$: Further Details

The solution to program $(3.1) - (3.3)$ yields:

$$c_L^t = y_L^t + \frac{\phi}{1-\phi}(y_H^t - c_H^t) \tag{A.1}$$

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

$$c_H^t = \frac{1-\phi}{\alpha} \left[\frac{1}{1+\gamma} \left(\frac{y_H^t}{w_H} \right)^{1+\gamma} + \alpha y_L^t + \frac{\alpha\phi}{1-\phi}y_H^t - \frac{1}{1+\gamma} \left(\frac{y_L^t}{w_H} \right)^{1+\gamma} \right] \tag{A.3}$$

$$y_H^t = \left[\frac{\lambda^t\phi(w_H)^{1+\gamma}}{(1-\pi)\phi + \theta_H^t} \right]^{\frac{1}{\gamma}} \tag{A.4}$$

where:

$$\lambda^t = \alpha \left[\pi(1-\phi) + (1-\pi)\phi \right] \tag{A.5}$$

is the multiplier on the government's budget constraint (3.2), and:

$$\theta_H^t = \frac{\lambda^t\phi}{\alpha} - (1-\pi)\phi \tag{A.6}$$

is the multiplier on the high-skill type's incentive-compatibility constraint (3.3).

Program (3.4) − (3.6): Further Details

The solution to program (3.4) − (3.6) yields:

$$\tilde{c}_L^t = \frac{\pi\left[(1-\phi)\tilde{y}_L^t + \phi\tilde{y}_H^t\right]}{\pi(1-\phi) + (1-\pi)\phi} \tag{A.7}$$

$$\tilde{y}_L^t = \left[\frac{\tilde{\lambda}^t(w_L)^{1+\gamma}}{\pi}\right]^{\frac{1}{\gamma}} \tag{A.8}$$

$$\tilde{c}_H^t = \frac{1-\pi}{\pi}\tilde{c}_L^t \tag{A.9}$$

$$\tilde{y}_H^t = \left[\frac{\tilde{\lambda}^t(w_H)^{1+\gamma}}{1-\pi}\right]^{\frac{1}{\gamma}} \tag{A.10}$$

where:

$$\tilde{\lambda}^t = \frac{\alpha\left[\pi(1-\phi)\pi + (1-\pi)\phi(1-\pi)\right]}{\pi(1-\phi) + (1-\pi)\phi} \tag{A.11}$$

is the multiplier on the government's budget constraint (3.5).

Program (3.7) − (3.9): Further Details

The solution to program (3.7) − (3.9) yields:

$$c_L^1 = y_L^1 + \frac{\phi}{1-\phi}(y_H^1 - c_H^1) \tag{A.12}$$

$$y_L^1 = \left[\frac{\lambda^1(1-\phi)(w_L)^{1+\gamma}(w_H)^{1+\gamma}}{\pi(1-\phi)(w_H)^{1+\gamma} - \theta_H(w_L)^{1+\gamma}}\right]^{\frac{1}{\gamma}} \tag{A.13}$$

$$c_H^1 = \frac{1-\phi}{\alpha}\left[\frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} + \alpha y_L^1 + \frac{\alpha\phi}{1-\phi}y_H^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} + \frac{\delta(1-p)}{1-\delta}\left[u_{LF}^t(\cdot) - u_{HF}^t(\cdot)\right]\right] \tag{A.14}$$

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

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

$$\lambda^1 = \alpha\left[\pi(1-\phi) + (1-\pi)\phi\right] \tag{A.16}$$

20

is the multiplier on the government's first-period budget constraint (3.8), and:

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

is the multiplier on the high-skill type's incentive-compatibility constraint (3.9).

8 Appendix B

Continuation Utility Functions: Further Details

After dropping the time superscripts, equations (5.3) and (5.4) can be solved to yield:

$$V_{iS} = \frac{u_{iS}(\cdot) + \delta(1 - q_S)V_{iF}(\cdot)}{1 - \delta q_S} \tag{B.1}$$

$$V_{iF} = \frac{(1 - \delta q_S)u_{iF}(\cdot) + \delta(1 - q_F)u_{iS}(\cdot)}{1 - \delta\left[\delta + (1 - \delta)(q_F + q_S)\right]} \tag{B.2}$$

Likewise, equations (5.5) and (5.6) can be solved to yield:

$$V_{MS} = \frac{u_{HS}(\cdot) + \delta(1 - q_S)V_{MF}(\cdot)}{1 - \delta q_S} \tag{B.3}$$

$$V_{MF} = \frac{(1 - \delta q_S)u_{LF}(\cdot) + \delta(1 - q_F)u_{HS}(\cdot)}{1 - \delta\left[\delta + (1 - \delta)(q_F + q_S)\right]} \tag{B.4}$$

Program (5.7) − (5.9): Further Details

The solution to program (5.7) − (5.9) yields:

$$c_L^1 = y_L^1 + \frac{\phi}{1 - \phi}(y_H^1 - c_H^1) \tag{B.5}$$

$$y_L^1 = \left[\frac{\lambda^1(1 - \phi)(w_L)^{1+\gamma}(w_H)^{1+\gamma}}{\pi(1 - \phi)(w_H)^{1+\gamma} - \theta_H(w_L)^{1+\gamma}}\right]^{\frac{1}{\gamma}} \tag{B.6}$$

$$c_H^1 = \frac{1 - \phi}{\alpha}\left[\begin{array}{c} \frac{1}{1+\gamma}\left(\frac{y_H^1}{w_H}\right)^{1+\gamma} + \alpha y_L^1 + \frac{\alpha\phi}{1-\phi}y_H^1 - \frac{1}{1+\gamma}\left(\frac{y_L^1}{w_H}\right)^{1+\gamma} \\ + \ p\delta\left[V_{MS}(\cdot) - V_{HS}(\cdot)\right] + (1 - p)\delta\left[V_{MF}(\cdot) - V_{HF}(\cdot)\right] \end{array}\right] \tag{B.7}$$

$$y_H^1 = \left[\frac{\lambda^1 \phi (w_H)^{1+\gamma}}{(1-\pi)\phi + \theta_H} \right]^{\frac{1}{\gamma}} \tag{B.8}$$

where:

$$\lambda^1 = \alpha \left[\pi(1-\phi) + (1-\pi)\phi \right] \tag{B.9}$$

is the multiplier on the government's first-period budget constraint (5.8), and:

$$\theta_H = \frac{\lambda^1 \phi}{\alpha} - (1-\pi)\phi \tag{B.10}$$

is the multiplier on the high-skill type's incentive-compatibility constraint (5.9).

References

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

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

[3] Auerbach, A. [2006], "Budget Windows, Sunsets, and Fiscal Control", *Journal of Public Economics*, 90, 87-100.

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

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

[6] Debortoli, D. and R. Nunes [2010], "Fiscal Policy under Loose Commitment", *Journal of Economic Theory*, 145, 1005-1032.

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

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

[9] 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.

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

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

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

[13] Guo, J-T. and A. Krause [2010b], "Dynamic Income Taxation without Commitment: Comparing Alternative Tax Systems", *University of York Discussion Papers in Economics*, No. 10-15.

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

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

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

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

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

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

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

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

TABLE 1

Baseline Parameter Values

π	0.55	γ	1.00	w_L	1.00
ϕ	0.25	δ	0.95	w_H	1.60
α	1.00	r	0.05	p	0.95

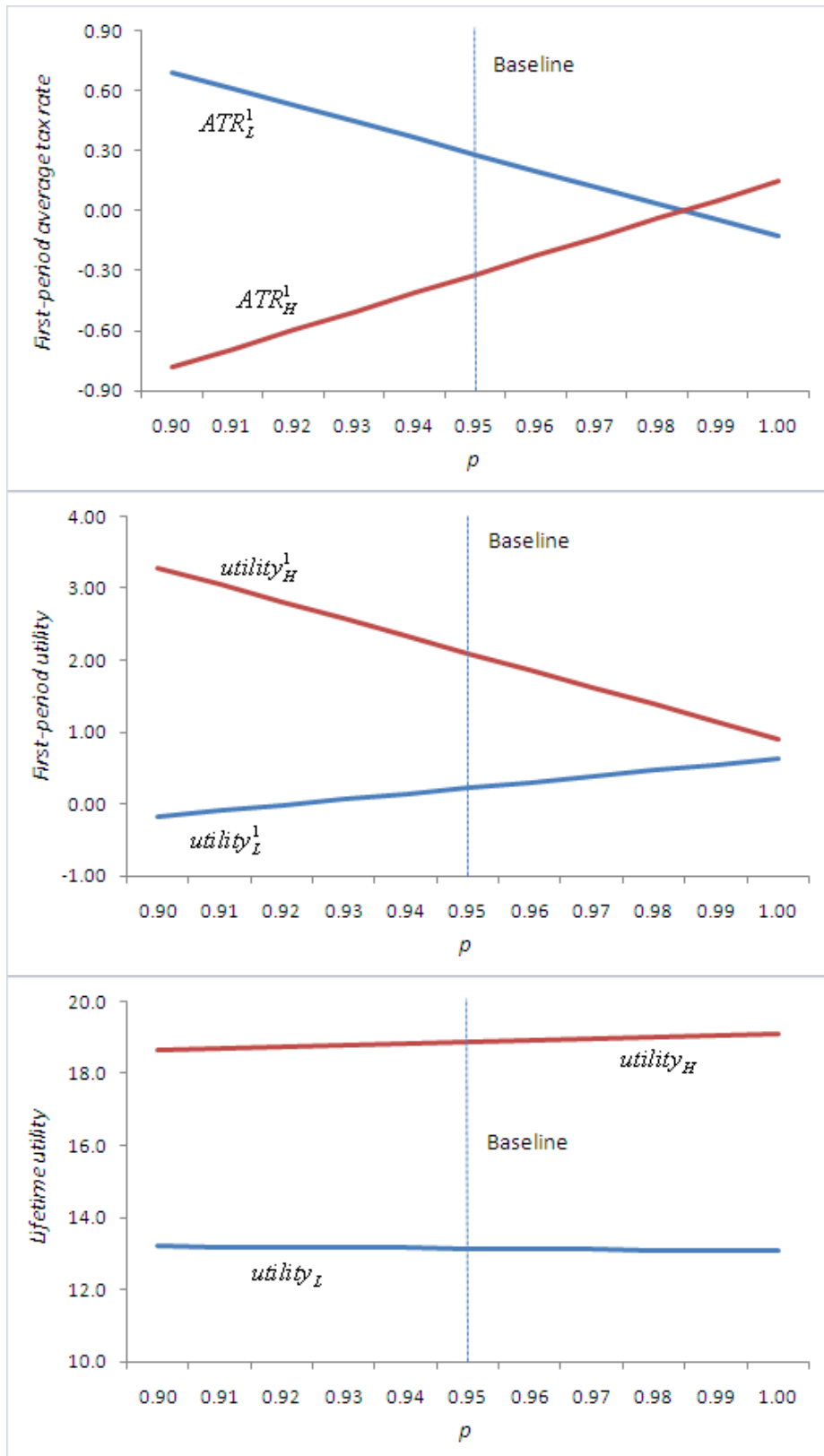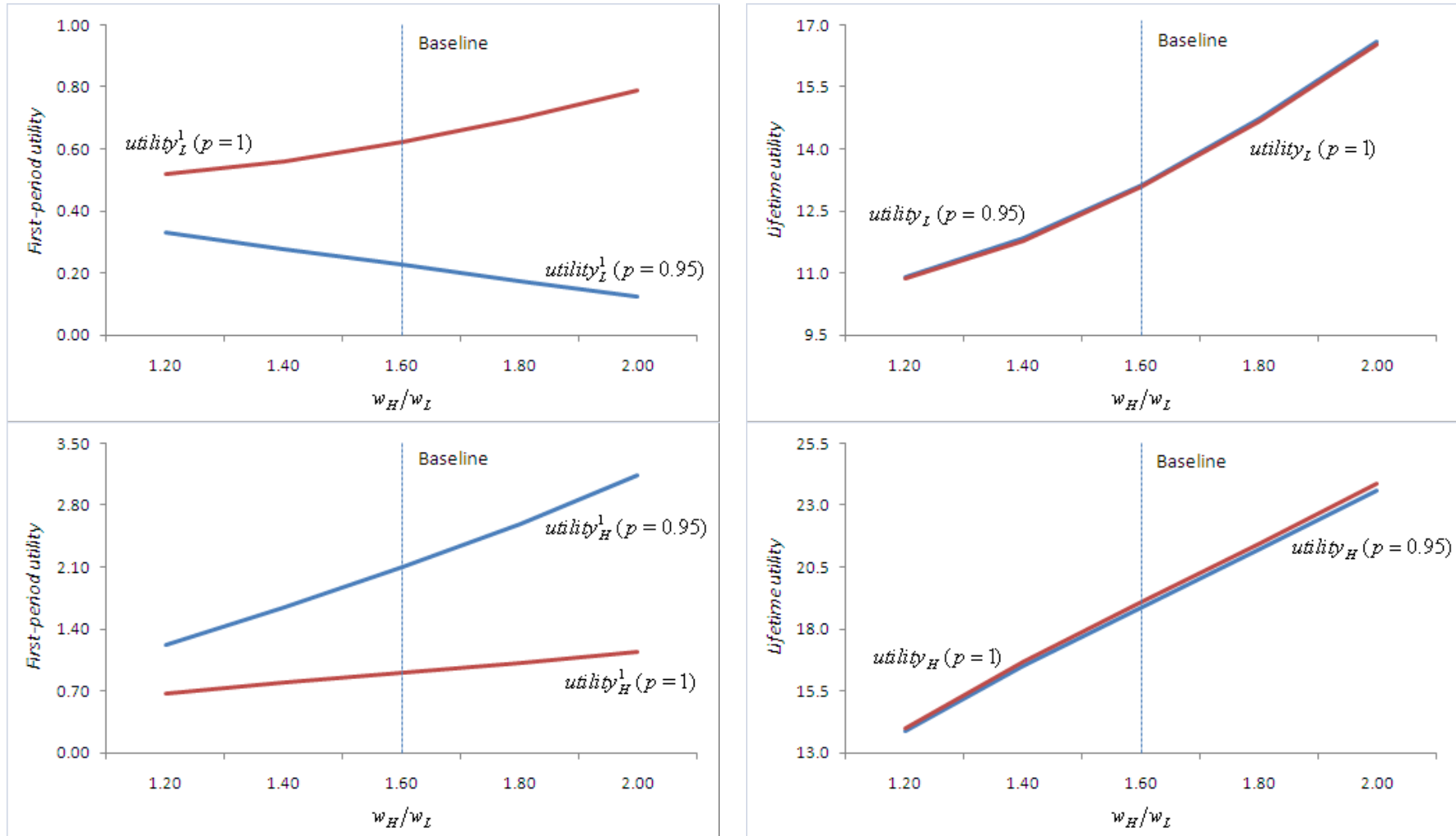|	*Period 1*	*Lifetime*
Utility: low-skill type	0.227	13.140
Utility: high-skill type	2.099	18.895

FIGURE 1

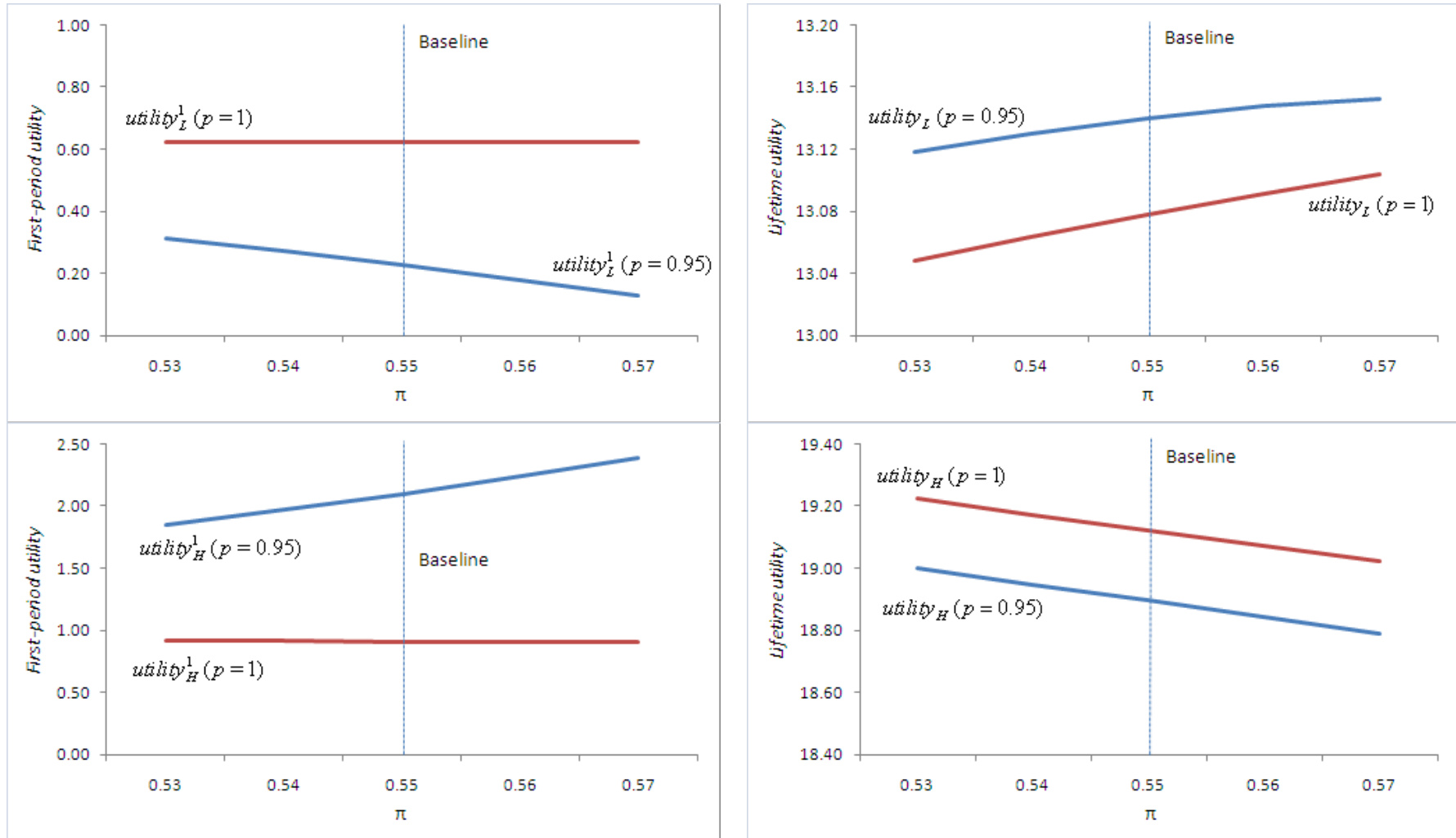## Effects of varying p

FIGURE 2

Effects of varying ϕ
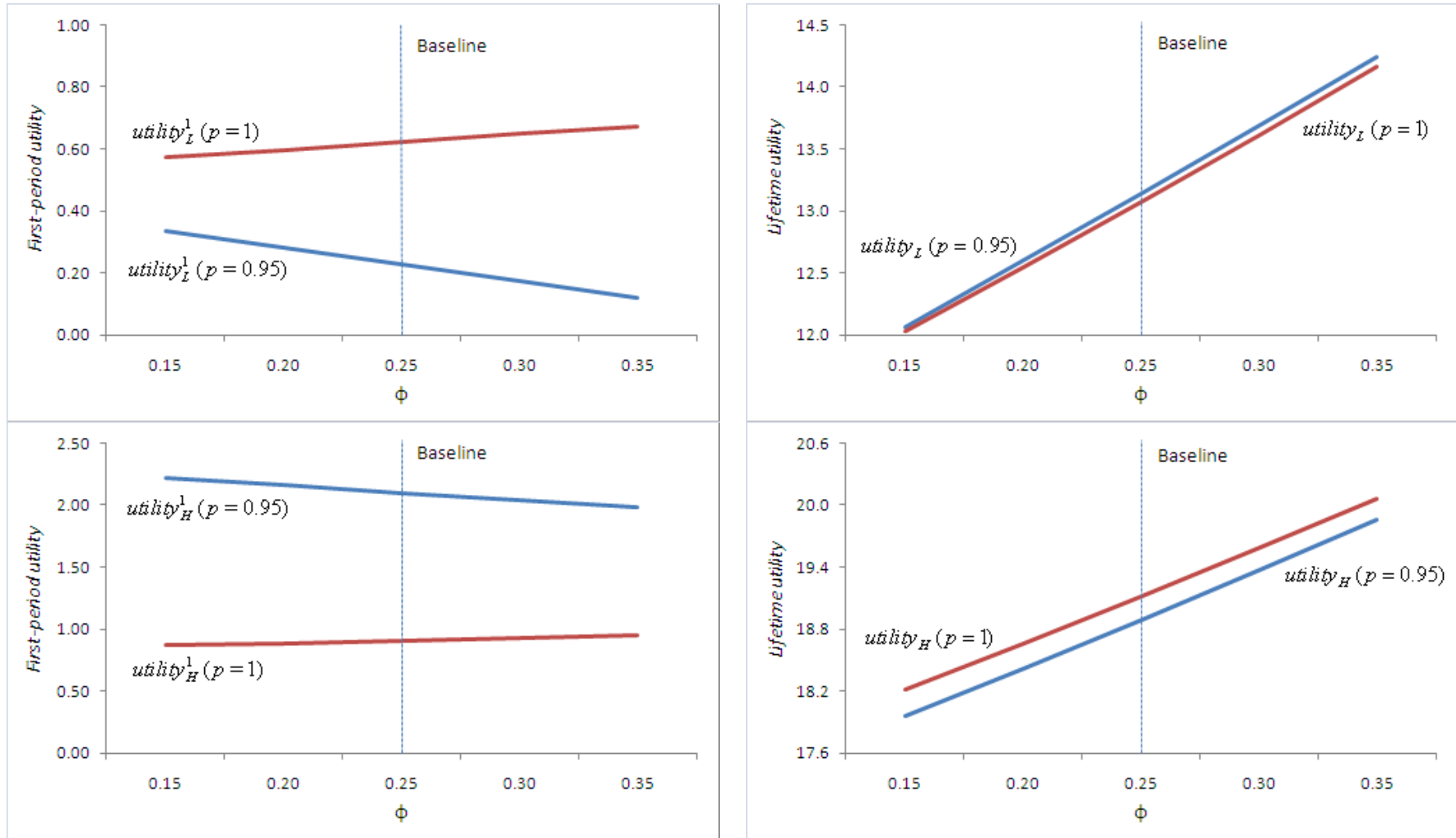
FIGURE 3

Effects of varying w_H/w_L

FIGURE 4

Effects of varying π

FIGURE 5

Effects of varying *r*