The Credibility of Commitment and Optimal Nonlinear Savings Taxation*

Yunmin Chen†
Shandong University
Jang-Ting Guo‡
University of California, Riverside
Alan Krause§
University of York

February 6, 2020

Abstract

Previous studies that examine the optimal nonlinear taxation of savings/capital have assumed either full-commitment or no-commitment by the government. This raises the question as to whether the results under full-commitment and no-commitment provide upper and lower bounds on the optimal marginal savings tax rates. We show that they do not. Specifically, we consider a model in which individuals attach some probability to whether or not the government can commit. When these probabilistic beliefs differ among individuals, optimal marginal savings tax rates may fall outside those under the polar cases of full-commitment and no-commitment.

Keywords: Savings taxation; commitment; multi-dimensional screening.

JEL Classifications: E60; H21; H24.

---

*This is a paper to which Alan Krause has made major contributions before he passed away. We also thank seminar participants at the University of York for helpful comments and suggestions. Any errors remain our responsibility.

†No. 27 ShandaNanlu, School of Economics, Shandong University, Jinan City, Shandong Province, P.R. China, 250100, E-mail: ymchen@sdu.edu.cn.

‡Corresponding author. Department of Economics, University of California, Riverside, CA, 92521, U.S.A., E-mail: guojt@ucr.edu.

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

Previous studies that examine the optimal nonlinear taxation of savings/capital have typically assumed full commitment by the government, while a few studies have considered no commitment. Under full-commitment, the government announces its tax policies for the present and future, and then simply implements those polices. Importantly, it is implicit that individuals completely believe that the government will implement its announced policies. Under no-commitment, the government re-optimizes its tax policies period by period, irrespective of any previous promises or announcements. Individuals are aware that the government will re-optimize each period. Therefore, under full-commitment, it is common knowledge among individuals and the government that the probability of commitment is one, whereas this probability is zero under no-commitment.

Our aim is to examine a setting that falls between the polar cases of full-commitment and no-commitment, and that better reflects realistic behavior and beliefs. To illustrate our thinking, consider for example corporate tax competition among governments. In order to entice a company to locate in its country, a government may promise and implement a low corporate tax rate. However, if the company does set-up operations in its country, the government will then be tempted to raise its corporate tax rate. This temptation follows because the company is now resident, and moving is costly. Moreover, when making its location decision, the company will be aware of this temptation and the possibility that it may face higher taxation in the future. The government will then be aware that the company is aware of this temptation, and so on. If one were to assume full-commitment, it is common knowledge that the government will never succumb to this temptation. On the other hand, it is common knowledge that the government will always succumb to this temptation under no-commitment. In our view, a more realistic setting is one we call ‘commitment without credibility’. In this case, the government sets taxes as it would under full-commitment, but it is aware that individuals attach some probability to re-optimization. Thus, the government’s promise to commit is not completely credible.

As in the related literature, we adopt the Mirrlees (1971) information-constrained ap-
proach to examine the optimal nonlinear taxation of labor income and savings. However, we do so under commitment-without-credibility and within an infinite-horizon overlapping generations (OLG) setting. In standard Mirrlees-style models in which individuals differ only by their skills, the existing literature has found that zero marginal savings taxation is optimal under full-commitment (based on Atkinson and Stiglitz 1976). But savings taxation will be progressive under no-commitment, in that the optimal marginal savings tax rates are increasing with respect to individuals’ skills (see, e.g., Brett and Weymark 2019; Farhi, et al. 2012). It is tempting to view the results under no-commitment as providing the upper-bound on the size of the optimal marginal savings tax distortions. We show that if individuals have the same beliefs regarding the probability of commitment, then optimal marginal savings tax rates do always fall between those under full-commitment and no-commitment. However, if individuals have different beliefs, optimal savings taxation is much more complicated. Our main result is that some individuals may face larger marginal savings tax distortions than they would under no-commitment, even though all individuals believe that there is some probability of commitment.

In summary, our analysis shows that a minor relaxation of the full-commitment assumption that individuals completely believe that the government can commit may yield substantive effects. Specifically, optimal nonlinear savings tax rates do not simply fall between those under full-commitment and no-commitment. The explanation for this counter-intuitive finding follows from the possibility that individuals may disagree vis-a-vis the probability of commitment. Under full-commitment and no-commitment, all individuals know and agree that the probability of commitment is one and zero, respectively. Therefore, individuals differ only by their skills. Under commitment-without-credibility, individuals differ by their skills as well as their beliefs. This makes the optimal tax problem more complicated, as it must now take into account this second source of heterogeneity. It will be shown that belief heterogeneity itself calls for mar-

\[1\] The first paper to examine Mirrlees-style taxation within an OLG setting is Ordover and Phelps (1979). There are now a number of papers that examine nonlinear taxation within OLG frameworks, though the specifics of our OLG model are most similar to Brett (2012) and Krause (2019).
ginal savings tax distortions. It is this new rationale for taxing savings that makes it possible for some individuals to face larger marginal savings tax distortions under commitment-without-credibility than they would under no-commitment.

The remainder of this paper is organized as follows. Section 2 outlines our modelling framework, Section 3 derives optimal nonlinear taxation under commitment-without-credibility, and Section 4 concludes. A number of mathematical derivations are contained in three appendices.

2 Preliminaries

We examine a simple model that makes the commitment-without-credibility optimal tax problem tractable. Despite this necessary simplicity, our model is sufficiently rich to capture the salient features of the problem. The model is an infinite-horizon overlapping generations (OLG) setting, in which each individual lives for two periods. Individuals work in the first period of their lives, and they may be a high-skill or low-skill worker, which creates a redistributive role for taxation. In the second period of their lives individuals are retired, and must live-off their savings. This creates a need for savings, as well as the possibility that the government cannot commit to its savings tax policy.

The skill levels of low-skill and high-skill workers are denoted by $a_1$ and $a_2$ respectively, with $a_2 > a_1 > 0$, so type 1 individuals are low-skill workers and type 2 individuals are high-skill workers. Each high-skill individual believes that the probability of commitment is high or low, denoted by $p^H$ and $p^L$ respectively, with $1 > p^H > p^L > 0$. Low-skill individuals have a common belief regarding the probability of commitment, but we do not specify its notation because it will be seen that their common belief plays no role in the design of the optimal tax system. Our model could be extended to incorporate heterogeneous beliefs among low-skill individuals. However, the optimal tax problem would become much more complicated, making it less clear as to which

---

2 The commitment-without-credibility optimal tax problem is a multi-dimensional screening problem. It is well-known that optimal tax problems with two unobserved characteristics can be difficult to solve, even in simple settings. See, for example, Boadway, et al. (2002) and the discussion within. Other recent papers that examine optimal taxation with multi-dimensional types include Jaquet and Lehman (2014) and Rothschild and Scheuer (2015).
incentive-compatibility constraints may bind (see Section 3). Let \( \pi_j \) denote the proportion of high-skill individuals who believe that the probability of commitment is \( j \), where \( j = L \) or \( H \) for low and high probability, respectively. We then have \( \pi_L + \pi_H = 1 \).

2.1 Consumers

All individuals live for two periods, period \( t \) and period \( t + 1 \). An individual born in period \( t \) consumes, works, and saves in period \( t \), and then lives-off their savings in period \( t + 1 \). Preferences are represented by the quasi-linear utility function:

\[
    u(c^t_{ij}) - l^t_{ij} + \delta u(x^{t+1}_{ij})
\]

where \( c^t_{ij} \) is type \( ij \)'s period \( t \) consumption, \( l^t_{ij} \) is type \( ij \)'s labor supply, and \( x^{t+1}_{ij} \) is type \( ij \)'s period \( t + 1 \) consumption. For high-skill individuals, \( i = 2 \) and \( j = L \) or \( H \) for low or high expected probability of commitment. For low-skill individuals the subscript \( j \) is redundant; thus we have \( c^t_i \), \( l^t_i \), and \( x^{t+1}_i \), where \( i = 1 \). The specification that the utility function is quasi-linear is stronger than the usual assumption in the literature that it is additively separable. However, quasi-linearity enables us to compare the size (not just sign) of the optimal marginal savings tax rates across different commitment regimes. The function \( u(\cdot) \) is increasing and strictly concave, while \( \delta \in (0, 1) \) denotes the discount factor. For future reference, we use \( y^t_{ij} = w^t a^t_{ij} l^t_{ij} \) to denote the pre-tax income of a type \( ij \) individual in period \( t \), where \( w^t \) is the wage rate and \( a^t_{ij} l^t_{ij} \) is type \( ij \)'s effective labor supply.

In the absence of taxation, individuals would choose \( c^t_{ij} \), \( l^t_{ij} \), \( s^t_{ij} \), and \( x^{t+1}_{ij} \) to maximize the utility function \( 2.1 \) subject to the budget constraints:

\[
    c^t_{ij} + s^t_{ij} \leq w^t a^t_{ij} l^t_{ij} \tag{2.2}
\]

\[
    x^{t+1}_{ij} \leq (1 + r^{t+1})s^t_{ij} \tag{2.3}
\]

where \( s^t_{ij} \) is type \( ij \)'s savings in period \( t \), and \( r^{t+1} \) is the rate-of-return on savings in period \( t + 1 \). It is shown in Appendix A that the solution to program \( 2.1 - 2.3 \) yields
the marginal condition:

\[ \frac{u'(c_{ij})}{\delta (1 + r^{t+1})u'(x_{ij}^{t+1})} = 1 \]  \hspace{1cm} (2.4)

However, in the presence of taxation, equation (2.4) might not be satisfied. The resulting marginal distortion is commonly interpreted as a ‘tax wedge’ or ‘implicit marginal tax rate’ on savings. Thus, we define:

\[ \tau_{ij}^t := 1 - \frac{u'(c_{ij})}{\delta (1 + r^{t+1})u'(x_{ij}^{t+1})} \]  \hspace{1cm} (2.5)

as the (implicit) marginal tax rate on savings in period \( t \) faced by a type \( ij \) individual.

2.2 Producer
The production side of our OLG model is standard. There is a representative profit-maximizing producer with a constant returns-to-scale production function:

\[ Q^t = F(K^t, Z^t) \]  \hspace{1cm} (2.6)

where \( Q^t \) is total output in period \( t \), \( K^t \) is the aggregate capital stock in period \( t \), and \( Z^t \) is total effective labor in period \( t \). Total effective labor consists of total low-skill and high-skill effective labor: \( Z^t = Z_1^t + Z_2^t \) where \( Z_1^t = N_1^t a_1 l_1^t \) and \( Z_2^t = \pi_L N_2^t a_2 l_2^L + \pi_H N_2^t a_2 l_2^H \), with \( N_1^t \) representing the number of type \( i \) workers in period \( t \). We make the simplifying assumption that \( N_1^t = N_2^t \), and that the populations of both low-skill and high-skill individuals grow at the same fixed rate of \( n \) per period, i.e., \( n = (N_1^{t+1} - N_1^t)/N_1^t \).

The production function can be rewritten as:

\[ Q^t = Z^t f(k^t) \]  \hspace{1cm} (2.7)

where \( k^t = K^t/Z^t \) is the ratio of capital to effective labor, and \( f(\cdot) \) is increasing and strictly concave. Profit maximization implies that:

\[ \frac{\partial Q^t}{\partial Z^t} = f(k^t) - f'(k^t)k^t = w^t \]  \hspace{1cm} (2.8)
\[ \frac{\partial Q^t}{\partial K^t} = f'(k^t) = r^t \]  

which are the standard conditions that the marginal product of (effective) labor equals the wage rate, and the marginal product of capital equals the rental rate.

### 2.3 Steady-State Equilibrium

We focus on the steady-state equilibrium of our OLG model, in which all variables per capita are constant through time. It is shown in Appendix A that the steady-state equilibrium condition is:

\[
f(k) = w \left[ c_1 + \pi_L c_{2L} + \pi_H c_{2H} + \frac{x_1 + \pi_L x_{2L} + \pi_H x_{2H}}{1+n} \right] + nk \tag{2.10}
\]

where the absence of the time superscript \( t \) represents the steady-state value of that variable.

### 3 Commitment Without Credibility

Under commitment-without-credibility, the government sets taxes as it would under full-commitment, but it is aware that individuals attach some probability to the government re-optimizing the savings tax after individuals have made their savings decisions. We examine optimal nonlinear taxation in steady-state, and under Mirrlees-style information constraints that the government cannot observe any individual’s skill type, nor their belief vis-a-vis the probability of commitment. Accordingly, the government’s choice of an optimal nonlinear tax schedule is equivalent to it choosing a steady-state allocation for each type of individual, \( c_1, y_1, x_1 \), \( c_{2L}, y_{2L}, x_{2L} \), \( c_{2H}, y_{2H}, x_{2H} \), and \( k \) to maximize:

\[
u(c_1) = \frac{y_1}{w} \delta u(x_1) + \sum_j \pi_j \left[ u(c_{2j}) - \frac{y_{2j}}{w} \delta u(x_{2j}) \right] \tag{3.1}
\]

subject to:

\[
f(k) \geq w \left[ c_1 + \pi_L c_{2L} + \pi_H c_{2H} + \frac{x_1 + \pi_L x_{2L} + \pi_H x_{2H}}{1+n} \right] + nk \tag{3.2}
\]
\begin{align}
\frac{u(c_{2H})}{w_2} - \frac{y_{2H}}{w_2} + \delta \left[ p^H u(x_{2H}) + (1 - p^H)u(x) \right] &\geq \frac{u(c_{2L})}{w_2} - \frac{y_{2L}}{w_2} + \delta \left[ p^H u(x_{2L}) + (1 - p^H)u(x) \right] \\
\frac{u(c_{2H})}{w_2} - \frac{y_{2H}}{w_2} + \delta \left[ p^H u(x_{2H}) + (1 - p^H)u(x) \right] &\geq \frac{u(c_1)}{w_2} - \frac{y_1}{w_2} + \delta \left[ p^H u(x_1) + (1 - p^H)u(x) \right] \\
\frac{u(c_{2L})}{w_2} - \frac{y_{2L}}{w_2} + \delta \left[ p^L u(x_{2L}) + (1 - p^L)u(x) \right] &\geq \frac{u(c_1)}{w_2} - \frac{y_1}{w_2} + \delta \left[ p^L u(x_1) + (1 - p^L)u(x) \right]
\end{align}
\begin{equation}
3.3
\end{equation}
\begin{equation}
3.4
\end{equation}
\begin{equation}
3.5
\end{equation}

where \( w = f(k) - f'(k)k \) by profit maximization (equation 2.8). Equation (3.1) is a utilitarian social welfare function,\(^3\) which is analogous to the social welfare function that the government would maximize under full-commitment. This reflects the government’s promise that it will not re-optimize the savings tax. Equation (3.2) is a restatement of the steady-state equilibrium condition (2.10), and it ensures that the chosen allocation is feasible. It also ensures, indirectly, that the government’s budget will be balanced.

Equations (3.3) – (3.5) are incentive-compatibility constraints. At this point we encounter the usual challenge associated with multi-dimensional screening problems, in that there are numerous (in our model, six) incentive-compatibility constraints that need to be considered.\(^4\) However, we analyze solutions to the optimal tax problem in which only (3.3) – (3.5) may be binding, for the following reasons. Consider incentive-compatibility constraint (3.3), which links the \( 2H \) and \( 2L \) types. Recall that \( 2H \) individuals are high-skill individuals who believe that there is a high probability of commitment. These individuals think that there is a good chance that the government will not re-optimize the savings tax, and therefore their retirement consumption will be \( x_{2H} \). But if the government does re-optimize, it will equalize consumption across all individuals as is optimal under no-commitment (see Appendix B). High-skill individuals would then have their savings taxed to subsidize the retirement consumption of low-skill individuals. Let \( x \) denote the equalized level of consumption in retirement. On the other hand, \( 2L \) individuals are high-skill individuals who believe that the probability of commitment is low. They think it is relatively more likely that the government will re-optimize the savings

\(^3\)Recall that there are equal numbers of low-skill and high-skill individuals.

\(^4\)If low-skill individuals were also distinguished by their beliefs regarding the probability of commitment, there would be twelve incentive-compatibility constraints to consider.
tax, and their retirement consumption will be only \( x \). Accordingly, \( 2L \) individuals will be relatively more reluctant to reveal their type, and they must be offered a generous allocation to do so. This means that \( 2H \) individuals will be attracted to the allocation intended for \( 2L \) individuals, but not vice versa. Incentive-compatibility constraint (3.3) prevents such mimicking behavior, by ensuring that the utility a \( 2H \) individual can expect from choosing \( \langle c_{2H}, y_{2H}, x_{2H} \rangle \) is greater than or equal to what they could obtain by choosing \( \langle c_{2L}, y_{2L}, x_{2L} \rangle \). Specifically, \( 2H \) individuals believe that the probability of commitment is \( p^H \), so with this probability they expect to be able to consume their savings in retirement, but with probability \( (1 - p^H) \) they expect that the government will re-optimize and their retirement consumption will be \( x \). Finally, incentive-compatibility constraints (3.4) and (3.5) reflect the standard assumption that the government seeks to redistribute from high-skill to low-skill individuals, which creates an incentive for each type of high-skill individual to mimic the low-skill, but not vice versa. In summary, for a broad range of plausible parameter values we believe that only incentive-compatibility constraints (3.3) – (3.5) will be binding. The numerical example below also confirms that only (3.3) – (3.5) are binding.

### 3.1 Homogeneous Beliefs

Before proceeding to our main results, we show that if individuals have the same beliefs regarding the probability of commitment, then optimal marginal savings tax rates must fall between those under full-commitment and no-commitment:

**Proposition 1** Suppose individuals have the same probabilistic beliefs, \( p^H = p^L = p \). The optimal marginal tax rates applicable to savings are: \( \tau^{NC}_2 > \tau_2 > 0 > \tau_1 > \tau^{NC}_1 \).

In Proposition 1, \( \tau^{NC}_2 \) and \( \tau^{NC}_1 \) represent the optimal marginal savings tax rates under no-commitment faced by high-skill and low-skill individuals, respectively. Under commitment-without-credibility, if individuals have the same probabilistic beliefs, the subscript \( j \) denoting belief type is redundant; thus we simply use \( \tau_2 \) and \( \tau_1 \) to represent the optimal marginal savings tax rates faced by high-skill and low-skill individuals, respectively. Proposition 1 confirms that without belief heterogeneity, the optimal marginal savings tax rates must fall between those under the polar cases of full-commitment (zero marginal savings taxation) and no-commitment (progressive marginal...
savings taxation). Indeed, if \( p = 1 \) (full-commitment) we obtain \( \tau_1 = \tau_2 = 0 \), and if \( p = 0 \) (no-commitment) we obtain \( \tau_2 = \tau_{2NC} > 0 \) and \( \tau_1 = \tau_{1NC} < 0 \). A full discussion of the intuition underlying the no-commitment results can be found in Brett and Weymark (2019) and Farhi, et al. (2012), but it may be summarized as follows.\(^5\)

High-skill individuals know that if the government re-optimizes the savings tax in their retirement period, it will redistribute some of their savings toward the low-skill. This creates an incentive for high-skill individuals to mimic low-skill individuals. Thus, in order to deter mimicking, the government brings forward consumption by high-skill individuals (by taxing their savings) and delays consumption by low-skill individuals (by subsidizing their savings). The intuition for the non-zero marginal savings tax rates under commitment-without-credibility is the same as under no-commitment, but they are closer to zero because there is some chance of commitment.

It is worth noting that the case of homogeneous beliefs considered here is similar to ‘loose commitment’ in the literature, the difference being that under loose commitment the government also explicitly recognizes that it may re-optimize with some probability. Thus, under loose commitment, it is common knowledge among individuals and the government that the probability of commitment is \( p \). It is straightforward to show that the pattern of optimal marginal savings tax rates in Proposition 1 also holds under loose commitment. Guo and Krause (2014) examine optimal nonlinear income taxation under loose commitment, but in a model without savings. Debortoli and Nunes (2010) examine capital and labor taxation under loose commitment in a representative-agent model.

3.2 Heterogeneous Beliefs

We now consider the general commitment-without-credibility setting in which \( p^H > p^L \), which leads to the following results:

**Proposition 2** Under commitment-without-credibility, the optimal marginal tax rates applicable to savings are: \( \tau_{2L} > \tau_{2H} > 0 > \tau_1 \).

High-skill individuals face positive marginal savings tax rates, while low-skill individuals face a negative marginal savings tax rate. This is qualitatively the same as under

\(^5\)In Brett and Weymark (2019) individuals work in both periods, while in Farhi, et al. (2012) individuals work only in the first period. In this respect, our analysis is closer to Farhi, et al. (2012).
no-commitment, and the intuition for the progressive pattern of optimal marginal savings taxation is analogous. Those high-skill individuals who believe that the probability of commitment is low (the 2L type) face the largest marginal savings tax distortion. This is because 2L individuals think it is relatively more likely that the government will re-optimize the savings tax. Accordingly, they must face the highest marginal savings tax rate to deter mimicking.

**Proposition 3** Type 2H individuals and low-skill individuals always face lower marginal savings tax rate distortions under commitment-without-credibility than under no-commitment, i.e., \( \tau_{2H} < \tau_{2NC} \) and \( \tau_{1NC} < \tau_1 \). However, type 2L individuals may face a greater marginal savings tax rate distortion under commitment-without-credibility than under no-commitment, i.e., \( \tau_{2L} > \tau_{2NC} \) is possible.

Type 2H individuals and low-skill individuals always face lower marginal savings tax rate distortions under commitment-without-credibility than under no-commitment. However, it is possible that 2L individuals face a larger marginal savings tax rate distortion under commitment-without-credibility than under no-commitment. This is our main result, and it may occur even though the only difference between full-commitment and commitment-without-credibility is that individuals under commitment-without-credibility attach some probability to re-optimization. The intuition is as follows. Under no-commitment individuals are distinguished only by their skills, as it is common knowledge that the government cannot commit. Under commitment-without-credibility, however, individuals are distinguished by their skills and their beliefs regarding the probability of commitment. Belief heterogeneity itself calls for marginal savings tax rate distortions. An increase in the degree of belief heterogeneity makes the tax contract intended for 2L individuals more attractive to 2H individuals. Accordingly, a higher marginal tax rate on savings by 2L individuals is required to deter mimicking. When the difference in beliefs is large enough, 2L individuals face a higher marginal savings tax rate under commitment-without-credibility than they would under no-commitment.

3.3 A Numerical Demonstration

In this subsection, we provide a numerical example to demonstrate the possibility stated in Proposition 3 that \( \tau_{2L} > \tau_{2NC} \). This possibility is demonstrated using empirically-
plausible parameter values, thus illustrating that it cannot be regarded as an unlikely occurrence. The numerical example also shows that only incentive-compatibility constraints (3.3)–(3.5) are binding. We assume that equal numbers of high-skill individuals believe that the probability of commitment is high or low, i.e., \( \pi_H = \pi_L = 1/2 \), and these probabilities are \( p^H = 0.9 \) and \( p^L = 0.1 \). That is, one-half of all high-skill individuals think that commitment is highly likely, while the other half think that commitment is highly unlikely.

The remainder of the calibration follows Krause (2019). We specify the utility function (2.1) as:

\[
\ln(c_{ij}) - l_{ij} + \delta \ln(x_{ij})
\]

which follows from Chetty (2006) who concludes that the coefficient of relative risk aversion is one (which implies log utility). Each period is assumed to be 20-years in length and the annual discount rate is assumed to be 2% (Kocherlakota 2010), making \( \delta = 0.67 \). The college wage premium is estimated to be approximately 60% (Fang 2006; Goldin and Katz 2007). Accordingly, we normalize \( a_1 = 1 \) and set \( a_2 = 1.6 \).

The production function (2.6) is assumed to be Cobb-Douglas:

\[
Q = K^{1/3} Z^{2/3}
\]

with the exponents (one-third and two-thirds) reflecting capital and labor shares of income in developed countries. Finally, we set \( n = 0.22 \), which reflects annual population growth of about 1%, and that each period of our model is 20-years in length.

The results of the numerical example are shown in Table 1. It can be seen that \( \tau_{2L} > \tau_{2H} > 0 > \tau_1 \), and \( 2H \) individuals and low-skill individuals face lower marginal savings tax rate distortions under commitment-without-credibility (CWC) than under no-commitment (NC). We also have \( \tau_{2L} > \tau_2^{NC} \), thus demonstrating the possibility identified in Proposition 3.

### 3.4 Discussion

Proposition 3 and the numerical example demonstrate the possibility that some individuals, in our model the \( 2L \) type, may face a higher marginal savings tax rate under
commitment-without-credibility than under no-commitment. This result is somewhat surprising, given that commitment-without-credibility is an intermediate setting between full-commitment and no-commitment. The result does, however, depend upon individuals having different beliefs regarding the probability of commitment. The question then arises as to why individuals may disagree vis-à-vis the probability of commitment. As assessing the credibility of a commitment promise is subjective, we simply assume that individuals may have different opinions regarding the credibility of that promise.

Another issue that arises is that regarding learning. In our model each individual lives for two periods, and observes only one commitment decision by the government. If individuals were able to observe many commitment decisions, it would be reasonable to expect that they would learn that the government’s commitment promise is credible, and update their beliefs accordingly. In that case, all individuals’ probabilistic beliefs would converge to one, i.e., the full-commitment setting. However, we think that the commitment-without-credibility setting remains a better description of observed reality than full-commitment or no-commitment.

Under commitment-without-credibility, the government promises to commit, but individuals do not completely believe that promise. This is because re-optimization yields a higher level of social welfare in the short run, so individuals recognize that the government will be tempted to re-optimize. Nevertheless, the government’s motivation in promising to commit is clear, as commitment yields the highest level of social welfare attainable in the long run. The opposite setting would be one in which the government states that it will re-optimize each period (no-commitment), but individuals believe that there is some chance that the government will not re-optimize. However, the motivations that the government and individuals have in this setting for their promises/beliefs are less clear, and we think that the commitment-without-credibility setting is a better description of observed reality.
4 Concluding Comments

The cases of full-commitment and no-commitment make strong assumptions regarding government behavior and the beliefs of individuals. Our specification of commitment-without-credibility represents an intermediate as well as a more realistic setting. We find that if individuals differ in their beliefs regarding the probability of commitment, some individuals may face larger marginal savings tax rate distortions than under no-commitment. This is possible even though all individuals believe that there is some probability of commitment. Therefore, the canonical result that zero marginal savings taxation is optimal under full-commitment (Atkinson and Stiglitz 1976) appears to be very sensitive to the assumption that commitment is completely credible.

5 Appendix A

A.1 Derivation of Equation (2.4)
The relevant first-order conditions from program (2.1) – (2.3) are:

\[ u'(c_{ij}^t) - \lambda^t = 0 \]  \hspace{1cm} (A.1)

\[ -\lambda^t + \lambda^{t+1}(1 + r^{t+1}) = 0 \]  \hspace{1cm} (A.2)

\[ \delta u'(x_{ij}^{t+1}) - \lambda^{t+1} = 0 \]  \hspace{1cm} (A.3)

where \( \lambda^t > 0 \) and \( \lambda^{t+1} > 0 \) are the multipliers on constraints (2.2) and (2.3), respectively. Algebraic manipulation of (A.1) – (A.3) yields equation (2.4).

A.2 Derivation of Equation (2.10)
In each period, equilibrium can be represented by the national accounting identity:

\[ F(K^t, Z^t) = C^t + I^t \]  \hspace{1cm} (A.4)

where \( C^t \) is aggregate consumption and \( I^t \) is aggregate investment in period \( t \). Equation
f(k^t) = N_1^t c_1^t + \pi_H N_2^t c_{2H}^t + \pi_L N_2^t c_{2L}^t + N_1^{t-1} x_1^t + \pi_H N_2^{t-1} x_{2H}^t + \pi_L N_2^{t-1} x_{2L}^t + K^{t+1} - K^t
\tag{A.5}

where \( N_1^t c_1^t + \pi_H N_2^t c_{2H}^t + \pi_L N_2^t c_{2L}^t \) is total consumption by young people in period \( t \), \( N_1^{t-1} x_1^t + \pi_H N_2^{t-1} x_{2H}^t + \pi_L N_2^{t-1} x_{2L}^t \) is total consumption by old people in period \( t \), and investment is equal to the change in the capital stock, \( K^{t+1} - K^t \), assuming for simplicity no capital depreciation. Dividing (A.5) by \( Z^t \) yields:

\[
f(k^t) = \frac{N_1^t c_1^t}{Z^t} + \frac{N_2^t (\pi_H c_{2H}^t + \pi_L c_{2L}^t)}{Z^t} + \frac{N_1^{t-1} x_1^t}{Z^t} + \frac{N_2^{t-1} (\pi_H x_{2H}^t + \pi_L x_{2L}^t)}{Z^t} + \frac{Z^{t+1}}{Z^t} k^{t+1} - k^t
\tag{A.6}

Under our assumptions that \( N_1^t = N_2^t \) and \( Z^t = N_1^t a_1 l_1^t + N_2^t a_2 (\pi_H l_{2H}^t + \pi_L l_{2L}^t) \), equation (A.6) becomes:

\[
f(k^t) = \frac{c_1^t + \pi_L c_{2L}^t + \pi_H c_{2H}^t + \frac{x_1^t + \pi_L x_{2L}^t + \pi_H x_{2H}^t}{1+n}}{a_1 l_1^t + \pi_L a_2 l_{2L}^t + \pi_H a_2 l_{2H}^t} + \frac{k^{t+1}(1+n)(a_1 l_1^{t+1} + \pi_L a_2 l_{2L}^{t+1} + \pi_H a_2 l_{2H}^{t+1})}{a_1 l_1^t + \pi_L a_2 l_{2L}^t + \pi_H a_2 l_{2H}^t} - k^t
\tag{A.7}

Given that \( y_{ij}^t = w^t a_i l_{ij}^t \), the steady-state version of (A.7) is:

\[
f(k) = \frac{w \left[ c_1 + \pi_L c_{2L} + \pi_H c_{2H} + \frac{x_1 + \pi_L x_{2L} + \pi_H x_{2H}}{1+n} \right]}{y_1 + \pi_L y_{2L} + \pi_H y_{2H}} + nk
\tag{A.8}

which is equation (2.10).

6 Appendix B

B.1 Optimal Nonlinear Taxation without Commitment

When it is common knowledge that the government cannot commit, all individuals know that the probability of commitment is zero, and therefore individuals differ only by their skills. As savings decisions are made in period \( t \), the stock of savings is fixed

\footnote{In addition to Brett and Weymark (2019) and Farhi, et al. (2012), there have been a number of recent papers that examine dynamic nonlinear taxation without commitment. See, e.g., Apps and Rees (2006), Berliant and Ledyard (2014), Krause (2009, 2017), Guo and Krause (2011, 2013, 2015a, 2015b),}
come period $t + 1$. Therefore, when the government re-optimizes in period $t + 1$, it will equate low-skill and high-skill consumption (Brett and Weymark 2019; Farhi, et al. 2012). Accordingly, the government’s behavior in steady-state can be described as follows. Choose a steady-state allocation $c_1, y_1, c_2, y_2, x$, and $k$ to maximize:

$$u(c_1) - \frac{y_1}{wa_1} + \delta u(x) + u(c_2) - \frac{y_2}{wa_2} + \delta u(x)$$  \hspace{1cm} (B.1)

subject to:

$$f(k) \geq \frac{w \left[ c_1 + c_2 + \frac{2x}{1+n} \right]}{y_1 + y_2} + nk$$  \hspace{1cm} (B.2)

$$u(c_2) - \frac{y_2}{wa_2} + \delta u(x) \geq u(c_1) - \frac{y_1}{wa_2} + \delta u(x)$$  \hspace{1cm} (B.3)

where $w = f(k) - f'(k)k$ by profit maximization (equation 2.8), and $x$ denotes the equalized level of consumption that both low-skill and high-skill individuals obtain in their retirement period. Equation (B.1) is a utilitarian social welfare function, equation (B.2) is the steady-state equilibrium condition, and equation (B.3) is the high-skill type’s incentive-compatibility constraint. As all individuals receive the same level of consumption (and hence utility) in their retirement period, the last term on both sides of equation (B.3) is redundant. But to assist interpretation of the incentive-compatibility constraint, we do not delete this term.

We first show that $r = n$. The first-order condition on $k$ can be written as:

$$\lambda (f'(k) - n) + \frac{\partial w}{\partial k} \left[ \frac{y_1}{w^2 a_1} - \frac{\theta_2 y_1}{w^2 a_1} + \frac{(1 + \theta_2) y_2}{w^2 a_2} \right] - \frac{\lambda}{(y_1 + y_2)} \left( c_1 + c_2 + \frac{2x}{1+n} \right) = 0$$  \hspace{1cm} (B.4)

where $\lambda > 0$ is the multiplier on constraint (B.2), and $\theta_2 > 0$ is the multiplier on constraint (B.3). The first-order conditions on $y_1$ and $y_2$ are:

$$-\frac{1}{wa_1} + \frac{\theta_2}{wa_2} + \frac{\lambda w}{(y_1 + y_2)^2} \left[ c_1 + c_2 + \frac{2x}{1+n} \right] = 0$$  \hspace{1cm} (B.5)

Aronsson and Sjogren (2016), and Morita (2016). However, the focus of those papers is not on savings taxation, but rather on how the government may use skill-type information revealed in earlier periods to implement first-best taxation in latter periods.
\[-\frac{(1 + \theta_2)}{w a_2} + \frac{\lambda w}{(y_1 + y_2)^2} \left[c_1 + c_2 + \frac{2x}{1 + n}\right] = 0 \]  
(B.6)

which together imply that:

\[\frac{y_1}{w^2 a_1} - \frac{\theta_2 y_1}{w^2 a_2} + \frac{(1 + \theta_2) y_2}{w^2 a_2} = \frac{\lambda}{(y_1 + y_2)} \left(c_1 + c_2 + \frac{2x}{1 + n}\right) \]  
(B.7)

Therefore, equation (B.4) reduces to \(f'(k) - n = 0\). Profit maximization implies that \(f'(k) = r\) (equation 2.9), which then leads to \(r = n\). For future reference, note also that (B.5) and (B.6) imply that \(\theta_2 = (a_2/a_1 - 1)/2\).

The first-order conditions on \(c_1, c_2,\) and \(x\) are, respectively:

\[(1 - \theta_2) u'(c_1) - \frac{\lambda w}{y_1 + y_2} = 0 \]  
(B.8)

\[(1 + \theta_2) u'(c_2) - \frac{\lambda w}{y_1 + y_2} = 0 \]  
(B.9)

\[2\delta u'(x) - \frac{2\lambda w}{(1 + n)(y_1 + y_2)} = 0 \]  
(B.10)

Algebraic manipulation of equations (B.8) and (B.10) yields:

\[1 - \frac{u'(c_1)}{\delta(1 + n)u'(x)} = \frac{-\theta_2}{1 - \theta_2} \]  
(B.11)

where \(1 - \theta_2 > 0\) by equation (B.8). Using equation (2.5) and \(r = n\), (B.11) implies that \(\tau_{1NC} < 0\). Likewise, algebraic manipulation of (B.9) and (B.10) yields:

\[1 - \frac{u'(c_2)}{\delta(1 + n)u'(x)} = \frac{\theta_2}{1 + \theta_2} \]  
(B.12)

Using equation (2.5) and \(r = n\), (B.12) implies that \(\tau_{2NC} > 0\).

7 Appendix C

C.1 Proof of Proposition 1

If \(p^H = p^L = p\), the commitment-without-credibility optimal tax problem becomes the
following: choose a steady-state allocation $c_1$, $y_1$, $x_1$, $c_2$, $y_2$, $x_2$, and $k$ to maximize:

$$u(c_1) - \frac{y_1}{wa_1} + \delta u(x_1) + u(c_2) - \frac{y_2}{wa_2} + \delta u(x_2)$$ (C.1)

subject to:

$$f(k) \geq \frac{w \left[ c_1 + c_2 + \frac{x_1 + x_2}{1+n} \right]}{y_1 + y_2} + nk$$ (C.2)

$$u(c_2) - \frac{y_2}{wa_2} + \delta \left[ pu(x_2) + (1-p)u(x) \right] \geq u(c_1) - \frac{y_1}{wa_2} + \delta \left[ pu(x_1) + (1-p)u(x) \right]$$ (C.3)

where equation (C.1) is the utilitarian social welfare function, equation (C.2) is the steady-state equilibrium condition, and equation (C.3) is the high-skill type’s incentive-compatibility constraint which reflects the common belief that the probability of commitment is $p$.

The first-order conditions on $y_1$ and $y_2$ can be written as:

$$\frac{-1}{wa_1} + \frac{\theta_2}{wa_2} + \frac{\lambda w}{(y_1 + y_2)^2} \left[ c_1 + c_2 + \frac{x_1 + x_2}{1+n} \right] = 0$$ (C.4)

$$\frac{-(1 + \theta_2)}{wa_2} + \frac{\lambda w}{(y_1 + y_2)^2} \left[ c_1 + c_2 + \frac{x_1 + x_2}{1+n} \right] = 0$$ (C.5)

where $\lambda > 0$ and $\theta_2 > 0$ are the multipliers on constraints (C.2) and (C.3), respectively. Solving equations (C.4) and (C.5) for $\theta_2$ yields $\theta_2 = (a_2/a_1 - 1)/2$, which is the same as under no-commitment. Moreover, as under no-commitment, it can be shown that $r = n$.

The first-order conditions on $c_1$, $x_1$, $c_2$, and $x_2$ are, respectively:

$$(1 - \theta_2)u'(c_1) - \frac{\lambda w}{y_1 + y_2} = 0$$ (C.6)

$$\delta (1 - \theta_2 p) u'(x_1) - \frac{\lambda w}{(1+n)(y_1 + y_2)} = 0$$ (C.7)

$$(1 + \theta_2)u'(c_2) - \frac{\lambda w}{y_1 + y_2} = 0$$ (C.8)

$$\delta (1 + \theta_2 p) u'(x_2) - \frac{\lambda w}{(1+n)(y_1 + y_2)} = 0$$ (C.9)
Algebraic manipulation of equations (C.6) and (C.7) yields:

\[ 1 - \frac{u'(c_1)}{\delta(1 + n)u'(x_1)} = -\theta_2(1 - p) \frac{1}{1 - \theta_2} \]  

(C.10)

where \(1 - \theta_2 > 0\) by equation (C.6). Comparing (B.11) and (C.10) establishes that \(\tau_1^{NC} < \tau_1 < 0\). Likewise, algebraic manipulation of (C.8) and (C.9) yields:

\[ 1 - \frac{u'(c_2)}{\delta(1 + n)u'(x_2)} = \frac{\theta_2(1 - p)}{1 + \theta_2} \]  

(C.11)

Comparing (B.12) and (C.11) establishes that \(\tau_2^{NC} > \tau_2 > 0\). ■

C.2 Proof of Proposition 2

We first show that \(r = n\) holds under commitment-without-credibility. The first-order condition on \(k\) in program (3.1) – (3.5) can be written as:

\[ \lambda (f'(k) - n) + \frac{\partial w}{\partial k} \left[ y_1 \left( \frac{1}{w^2 a_1} - \frac{\theta_2 H + \theta_2 L}{w^2 a_2} \right) + \frac{y_2 H (\pi_H + \gamma_2 H + \theta_2 H)}{w^2 a_2} + \frac{y_2 L (\pi_L - \gamma_2 H + \theta_2 L)}{w^2 a_2} \right] \]

\[ - \frac{\partial w}{\partial k} \lambda \left[ c_1 + \pi_L c_2 L + \pi_H c_2 H + \frac{x_1 + \pi_L x_2 L + \pi_H x_2 H}{1 + n} \right] = 0 \]  

(C.12)

where \(\lambda > 0, \gamma_2 H > 0, \theta_2 H > 0, \) and \(\theta_2 L > 0\) are the multipliers on constraints (3.2) – (3.5), respectively. The first-order conditions on \(y_1, y_2 H, \) and \(y_2 L\) are:

\[ -\frac{1}{wa_1} + \frac{\theta_2 H + \theta_2 L}{wa_2} + \frac{\lambda w \left[ c_1 + \pi_L c_2 L + \pi_H c_2 H + \frac{x_1 + \pi_L x_2 L + \pi_H x_2 H}{1 + n} \right]}{\left( y_1 + \pi_L y_2 L + \pi_H y_2 H \right)^2} = 0 \]  

(C.13)

\[ -\left( \pi_H + \gamma_2 H + \theta_2 H \right) \frac{wa_2}{wa_2} + \frac{\pi_H \lambda w \left[ c_1 + \pi_L c_2 L + \pi_H c_2 H + \frac{x_1 + \pi_L x_2 L + \pi_H x_2 H}{1 + n} \right]}{\left( y_1 + \pi_L y_2 L + \pi_H y_2 H \right)^2} = 0 \]  

(C.14)

\[ -\left( \pi_L - \gamma_2 H + \theta_2 L \right) \frac{wa_2}{wa_2} + \frac{\pi_L \lambda w \left[ c_1 + \pi_L c_2 L + \pi_H c_2 H + \frac{x_1 + \pi_L x_2 L + \pi_H x_2 H}{1 + n} \right]}{\left( y_1 + \pi_L y_2 L + \pi_H y_2 H \right)^2} = 0 \]  

(C.15)

Using (C.13) – (C.15), equation (C.12) reduces to \(f'(k) - n = 0\), which by profit maximization (equation 2.9) establishes that \(r = n\).
The first-order conditions on \( c_1 \) and \( x_1 \) are, respectively:

\[
(1 - \theta_{2H} - \theta_{2L})u'(c_1) - \frac{\lambda w}{y_1 + \pi_L y_2L + \pi_H y_{2H}} = 0 \tag{C.16}
\]

\[
\delta(1 - \theta_{2H}p^H - \theta_{2L}p^L)u'(x_1) - \frac{\lambda w}{(1 + n)(y_1 + \pi_L y_2L + \pi_H y_{2H})} = 0 \tag{C.17}
\]

Algebraic manipulation of equations (C.16) and (C.17) yields:

\[
1 - \frac{u'(c_1)}{\delta(1 + n)u'(x_1)} = -\frac{[\theta_{2H}(1 - p^H) + \theta_{2L}(1 - p^L)]}{1 - \theta_{2H} - \theta_{2L}} \tag{C.18}
\]

where \( 1 - \theta_{2H} - \theta_{2L} > 0 \) by equation (C.16). Using equation (2.5) and \( r = n \), equation (C.18) implies that \( \tau_1 < 0 \).

The first-order conditions on \( c_{2H} \) and \( x_{2H} \) are, respectively:

\[
(\pi_H + \gamma_{2H} + \theta_{2H})u'(c_{2H}) - \frac{\pi_H \lambda w}{y_1 + \pi_L y_2L + \pi_H y_{2H}} = 0 \tag{C.19}
\]

\[
\delta(\pi_H + \gamma_{2H}p^H + \theta_{2L}p^L)u'(x_{2H}) - \frac{\pi_H \lambda w}{(1 + n)(y_1 + \pi_L y_2L + \pi_H y_{2H})} = 0 \tag{C.20}
\]

Algebraic manipulation of equations (C.19) and (C.20) yields:

\[
1 - \frac{u'(c_{2H})}{\delta(1 + n)u'(x_{2H})} = \frac{(1 - p^H)(\gamma_{2H} + \theta_{2H})}{\pi_H + \gamma_{2H} + \theta_{2H}} \tag{C.21}
\]

Using equation (2.5) and \( r = n \), equation (C.21) implies that \( \tau_{2H} > 0 \).

The first-order conditions on \( c_{2L} \) and \( x_{2L} \) are, respectively:

\[
(\pi_L - \gamma_{2H} + \theta_{2L})u'(c_{2L}) - \frac{\pi_L \lambda w}{y_1 + \pi_L y_2L + \pi_H y_{2H}} = 0 \tag{C.22}
\]

\[
\delta(\pi_L - \gamma_{2H}p^H + \theta_{2L}p^L)u'(x_{2L}) - \frac{\pi_L \lambda w}{(1 + n)(y_1 + \pi_L y_2L + \pi_H y_{2H})} = 0 \tag{C.23}
\]

Algebraic manipulation of equations (C.22) and (C.23) yields:

\[
1 - \frac{u'(c_{2L})}{\delta(1 + n)u'(x_{2L})} = \frac{\theta_{2L}(1 - p^L) - \gamma_{2H}(1 - p^H)}{\pi_L - \gamma_{2H} + \theta_{2L}} \tag{C.24}
\]
where $\pi_L - \gamma_{2H} + \theta_{2L} > 0$ by equation (C.22). Recall that $p^H > p^L$. Also, using equations (C.14) and (C.15) we obtain:

$$\frac{\theta_{2L} - \gamma_{2H}}{\pi_L} = \frac{\theta_{2H} + \gamma_{2H}}{\pi_H}$$

which establishes that $\theta_{2L} > \gamma_{2H}$. Using equation (2.5) and $r = n$, equation (C.24) now implies that $\tau_{2L} > 0$.

Finally, using (C.21) and (C.24) we obtain:

$$\tau_{2L} - \tau_{2H} = \frac{\theta_{2L}(1 - p^L) - \gamma_{2H}(1 - p^H)}{\pi_L - \gamma_{2H} + \theta_{2L}} - \frac{(1 - p^H)(\gamma_{2H} + \theta_{2H})}{\pi_H + \gamma_{2H} + \theta_{2H}}$$

which can be manipulated to yield:

$$\tau_{2L} - \tau_{2H} = \frac{(p^H - p^L) [(\pi_L + \theta_{2L})(\gamma_{2H} + \theta_{2H}) + \pi_H \gamma_{2H}]}{(\pi_L - \gamma_{2H} + \theta_{2L})(\pi_H + \gamma_{2H} + \theta_{2H})}$$

which establishes that $\tau_{2L} > \tau_{2H}$. ■

**C.3 Proof of Proposition 3**

Using (B.11) and (C.18) we obtain:

$$\tau_1 - \tau_1^{NC} = - \frac{[\theta_{2H}(1 - p^H) + \theta_{2L}(1 - p^L)]}{1 - \theta_{2H} - \theta_{2L}} + \frac{\theta_2}{1 - \theta_2}$$

which can be manipulated to yield:

$$\tau_1 - \tau_1^{NC} = \frac{(1 - \theta_2) [\theta_{2H}p^H + \theta_{2L}p^L] + \theta_2 - \theta_{2H} - \theta_{2L}}{(1 - \theta_{2H} - \theta_{2L})(1 - \theta_2)}$$

Using (C.13) and (C.14) we obtain:

$$\frac{1}{wa_1} - \frac{1}{wa_2} - \frac{\theta_{2H} + \theta_{2L}}{wa_2} = \frac{\theta_{2H} + \gamma_{2H}}{\pi_H wa_2}$$

which leads to:

$$\frac{1}{2} \left( \frac{a_2}{a_1} - 1 \right) - \frac{\theta_{2H} + \theta_{2L}}{2} = \frac{\theta_{2H} + \gamma_{2H}}{2\pi_H}$$
Recall that $\theta_2 = (a_2/a_1 - 1)/2$. Therefore, equation (C.31) becomes:

$$\theta_2 - \theta_{2H} - \theta_{2L} = \frac{\theta_{2H} + \gamma_{2H}}{2\pi_H} - \frac{\theta_{2H} + \theta_{2L}}{2} \quad (C.32)$$

Algebraic manipulation reveals that the right-hand side of equation (C.32) equals zero if and only if:

$$\theta_{2H}(1 - \pi_H) = \pi_H \theta_{2L} - \gamma_{2H} \quad (C.33)$$

From (C.25) we obtain:

$$\pi_H \theta_{2L} = (\theta_{2H} + \gamma_{2H}) \pi_L + \gamma_{2H} \pi_H \quad (C.34)$$

Thus, after algebraic manipulation:

$$\theta_{2H}(1 - \pi_H) = \pi_H \theta_{2L} - \gamma_{2H} \quad \text{iff} \quad \pi_H + \pi_L = 1 \quad (C.35)$$

which is true. Therefore, $\theta_2 - \theta_{2H} - \theta_{2L} = 0$ and equation (C.29) reduces to:

$$\tau_1 - \tau_1^{NC} = \frac{(1 - \theta_2) [\theta_{2H}\pi^H + \theta_{2L}\pi^L]}{(1 - \theta_{2H} - \theta_{2L})(1 - \theta_2)} \quad (C.36)$$

which establishes that $\tau_1 > \tau_1^{NC}$.

Using (B.12) and (C.21) we obtain:

$$\tau_2^{NC} - \tau_{2H} = \frac{\theta_2}{1 + \theta_2} - \frac{(1 - p_H)(\gamma_{2H} + \theta_{2H})}{\pi_H + \gamma_{2H} + \theta_{2H}} \quad (C.37)$$

which can be manipulated to yield:

$$\tau_2^{NC} - \tau_{2H} = \frac{p_H (1 + \theta_2) (\gamma_{2H} + \theta_{2H}) + \theta_2 \pi_H - \gamma_{2H} - \theta_{2H}}{(1 + \theta_2) (\pi_H + \gamma_{2H} + \theta_{2H})} \quad (C.38)$$

From (C.31) we obtain:

$$\theta_2 \pi_H - \frac{\pi_H(\theta_{2H} + \theta_{2L})}{2} = \frac{\theta_{2H} + \gamma_{2H}}{2} \quad (C.39)$$
which can be rewritten as:

\[ \theta_{2H} - \gamma_{2H} - \theta_{2H} = \frac{\pi_H(\theta_{2H} + \theta_{2L})}{2} - \frac{\theta_{2H} + \gamma_{2H}}{2} \]  

(C.40)

Algebraic manipulation, along with use of (C.34), reveals that the right-hand side of equation (C.40) equals zero if and only if:

\[ (\pi_H + \pi_L) \theta_{2H} + \gamma_{2H} = \theta_{2H} + \gamma_{2H} \]  

(C.41)

which is true given that \( \pi_H + \pi_L = 1 \). Therefore, equation (C.38) reduces to:

\[ \tau_2^{NC} - \tau_{2H} = \frac{p^H (1 + \theta_2) \gamma_{2H}}{(1 + \theta_2) (\pi_H + \gamma_{2H} + \theta_{2H})} \]  

(C.42)

which establishes that \( \tau_2^{NC} > \tau_{2H} \).

Using (B.12) and (C.24) we obtain:

\[ \tau_{2L} - \tau_2^{NC} = \frac{\theta_{2L}(1 - p^L) - \gamma_{2H}(1 - p^H)}{\pi_L - \gamma_{2H} + \theta_{2L}} - \frac{\theta_2}{1 + \theta_2} \]  

(C.43)

which can be manipulated to yield:

\[ \tau_{2L} - \tau_2^{NC} = \frac{\theta_{2L}(1 - p^L) - \gamma_{2H}(1 - p^H) - \theta_2 (\pi_L - \gamma_{2H} p^H + \theta_{2L} p^L)}{(\pi_L - \gamma_{2H} + \theta_{2L}) (1 + \theta_2)} \]  

(C.44)

Algebraic manipulation of (C.44) reveals that \( \tau_{2L} > \tau_2^{NC} \) if and only if \( p^H / p^L > \theta_{2L} / \gamma_{2H} \).

Table 1 provides an example where \( \tau_{2L} > \tau_2^{NC} \).
References


TABLE 1
A Numerical Example*

<table>
<thead>
<tr>
<th>Parameter values</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_H = 0.5$</td>
<td>$\delta = 0.67$</td>
</tr>
<tr>
<td>$\pi_L = 0.5$</td>
<td>$n = 0.22$</td>
</tr>
<tr>
<td>$p_H = 0.9$</td>
<td>$p_L = 0.1$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CWC</th>
<th>NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{2H}$</td>
<td>0.023</td>
<td>0.231</td>
</tr>
<tr>
<td>$\tau_{2L}$</td>
<td>0.300</td>
<td>0.231</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>−0.300</td>
<td>−0.429</td>
</tr>
</tbody>
</table>

*Under CWC, incentive-compatibility constraints (3.3), (3.4), and (3.5) are all binding. All omitted incentive-compatibility constraints are slack.