

Essays on Optimal Taxation

by

Catarina Reis

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

© Catarina Reis, MMVII. All rights reserved.

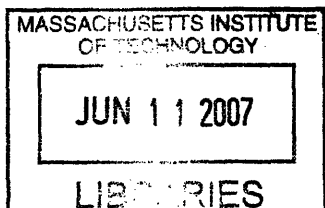
The author hereby grants to MIT permission to reproduce and
distribute publicly paper and electronic copies of this thesis document
in whole or in part.

Author
Department of Economics
May 15, 2007

Certified by
Peter Diamond
Institute Professor
Thesis Supervisor

Certified by
Michael Golosov
Assistant Professor
Thesis Supervisor

Accepted by
Peter Temin, Elisha Gray II Professor of Economics
Chairman, Department Committee on Graduate Students



ARCHIVES

Essays on Optimal Taxation

by

Catarina Reis

Submitted to the Department of Economics
on May 15, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

This thesis studies the optimal income tax scheme in four different settings. Chapter 1 focuses on the implications of lack of commitment for the optimal labor and capital income tax rates. It finds that it is optimal to converge to zero capital income taxes and positive labor income taxes in the long run. The government will follow the optimal plan as long as its debt is low enough, which implies that the lack of commitment may lead to some asset accumulation in the short run. Chapter 2 determines the optimal tax schedule when education is endogenous and observable, in a setting where agents have heterogeneous abilities. It finds that, for each ability level, it is optimal to subsidize monetary educational costs at the same marginal rate at which income is being taxed. Chapter 3 finds that when entrepreneurial labor income cannot be observed separately from capital income, then it is optimal to have positive capital taxation in the long run. Chapter 4 finds that if human capital expenses are unobservable, then in the optimal plan human capital accumulation will be distorted in the long run.

Thesis Supervisor: Peter Diamond
Title: Institute Professor

Thesis Supervisor: Michael Golosov
Title: Assistant Professor

Acknowledgments

I am grateful to my advisors Peter Diamond and Mike Golosov for their guidance and support. I would also like to thank Marios Angeletos, Ricardo Caballero, and Ivan Werning for helpful comments and suggestions. I am grateful to Fundacao para a Ciencia e Tecnologia, the National Science Foundation, and the Department of Economics at MIT for financial support.

Contents

1	Taxation without Commitment	13
1.1	Introduction	13
1.2	Taxation with Commitment	18
1.2.1	Model Setup	18
1.2.2	Allocations Attainable under Commitment	20
1.2.3	Optimal Taxes with Commitment	22
1.3	Sustainable Equilibria without Commitment	24
1.3.1	Game Setup	24
1.3.2	Strategies	25
1.3.3	Sustainable Equilibria	26
1.3.4	Worst Sustainable Equilibrium	28
1.3.5	Characterization of Sustainable Outcomes	30
1.4	Best Sustainable Equilibrium	31
1.4.1	Sequence Approach	31
1.4.2	Recursive Approach	32
1.4.3	Steady State	34
1.5	Numerical Simulations	37
1.5.1	Short Run Dynamics	37
1.5.2	Long Run Implications	40
1.6	Concluding Remarks	42
1.7	Appendix	42
1.7.1	Proof of Lemma 1	42

1.7.2	Proof of Lemma 2	45
1.7.3	Proof of Lemma 3	46
1.7.4	Proof of Lemma 4	47
1.7.5	Equivalence of Sequence and Recursive Approaches	48
1.7.6	Impossibility of non zero Capital Taxes in Steady State	50
2	Education and Optimal Taxation	55
2.1	Introduction	55
2.2	The model	58
2.2.1	Time Cost of Education	60
2.2.2	Monetary Cost of Education	62
2.2.3	Time and Monetary Cost of Education	65
2.3	Comparing Different Models of Education	65
2.3.1	Unobservable Education	66
2.3.2	Myopic Social Planner	68
2.3.3	Simulation of Optimal Taxes for Different Models of Education	69
2.3.4	Actual Taxes in the US	73
2.4	Concluding Remarks	75
2.5	Appendix	76
2.5.1	Derivation of the Optimal Tax First-Order Conditions	76
3	Taxation of Entrepreneurial Capital in a Deterministic Setting	79
3.1	Introduction	79
3.2	Model Setup	80
3.2.1	Households	81
3.2.2	Firms and Market Clearing	82
3.3	Optimal Plan	83
3.4	An Example	86
3.5	Concluding Remarks	87
3.6	Appendix	88
3.6.1	Optimal Taxation with Two Taxable Labor Inputs	88

3.6.2	Optimal Taxation with One Untaxable Labor Input	88
3.6.3	Optimal Joint Capital and Labor Tax	90
4	Optimal Taxation with Unobservable Investment in Human Capital	93
4.1	Introduction	93
4.2	Benchmark Model	94
4.2.1	Households	94
4.2.2	Firms and Market Clearing	96
4.2.3	Optimal Plan	96
4.3	Unobservable Investment in Human Capital	98
4.4	Concluding Remarks	100
4.5	Appendix	101
4.5.1	Observable Investment in Human Capital	102
4.5.2	Unobservable Investment in Human Capital	103
4.5.3	Taxation of Physical Capital	103

List of Figures

1-1	Transition Dynamics with and without Commitment	39
1-2	Steady States achievable with and without Commitment	41
2-1	Estimated Density and Cumulative Distribution of Abilities	70
2-2	Estimated Ratios $\frac{1-F(\theta)}{f(\theta)}$ in grey and $\frac{1-F(\theta)}{\theta f(\theta)}$ in black	71
2-3	Estimated Optimal Tax Schedule	72
2-4	Expected Wages and Taxes for each Schooling Level	74

Chapter 1

Taxation without Commitment

This chapter considers a Ramsey model of linear capital and labor income taxation in which a benevolent government cannot commit ex-ante to a sequence of taxes for the future. In this setup, if the government is allowed to borrow and lend to the consumers, the optimal capital income tax is zero in the long run. This result stands in marked contrast with the recent literature on optimal taxation without commitment, which imposes budget balance and typically finds that the optimal capital income tax does not converge to zero. Since it is efficient to backload incentives, breaking budget balance allows the government to generate surplus that reduces its debt or increases its assets over time until the lack of commitment is no longer binding and the economy is back in the full commitment solution. Therefore, while the lack of commitment does not change the optimal capital tax in the long run, it may impose an upper bound on the level of long run debt.

1.1 Introduction

This chapter explores the issue of optimal capital and labor income taxation when the government cannot commit to future taxes. By allowing the government to borrow and lend from households, the model generates results substantially different from the ones found by the previous literature on taxation without commitment. The reason for this is that governments with more assets need to use less distortionary taxation,

which means that the deadweight burden of additional revenue is smaller. Thus, the incentive to default can be reduced by allowing asset accumulation.

A traditional question in the optimal taxation literature concerns the extent to which capital taxes should be used to finance public spending. While in the short run it is optimal to tax capital to collect costless revenue from a sunk investment, in the long run using this source of taxation will distort the accumulation of capital. Chamley (1986) and Judd (1985) show that in an economy of infinitely lived agents capital taxes lead to intertemporal distortions that compound over time, creating an infinite wedge between marginal utility in different periods. Therefore, in the long run, the capital income tax should asymptote to zero.

It has been believed the result of zero capital taxes in the long run critically hinges on the ability of the government to commit ex-ante to a sequence of future taxes. Namely, Judd (1985) says that his "results indicate that redistribution of income through capital income taxation is effective only if it is unanticipated and will persist only if policy-makers cannot commit themselves to low taxation in the long run." Later work by Benhabib and Rustichini (1997) and Phelan and Stachetti (2001) confirms this intuition by finding that when commitment binds, the long run capital tax will not be zero. Using numerical simulations, Fernandez-Villaverde and Tsyvinsky (2002) find that, in general, commitment will bind if the government is impatient enough, since the future reward of a better equilibrium will not be enough to prevent the government from deviating from the predefined plan. However, all these papers assume that the government has to keep budget balance in each period.

This chapter shows that if instead the government is allowed to borrow and lend to consumers, the optimal capital tax still converges to zero in the long run, as in the full commitment case. The reason for this is that a government with a large amount of assets will not have an incentive to default since it does not need to use much distortionary taxation to finance its spending. Thus, governments can use asset accumulation (or debt reduction) as a commitment device for the future. As long as commitment binds, there will be an incentive to increase the government's assets. This is consistent with the result found independently by Dominguez (2006), who

analyzes the model in Benhabib and Rustichini (1997) for the case where bonds are allowed but the value of default is exogenous and depends only on capital.

Although the economy without commitment converges to a steady state where commitment does not bind and capital income taxes are zero, some steady states that were feasible in the economy with commitment will never be reached without commitment. If the economy with commitment converged to a steady state with high government debt that is no longer sustainable without commitment, then in the economy without commitment the government will have to accumulate more assets in the short run and will converge to a new steady state with lower debt. Hence, while the lack of commitment does not change the optimal capital tax in the long run, it may impose an upper bound on the long run level of debt.

A rather unexpected consequence of the lack of commitment is that capital levels will tend to be higher in the long run when there is no commitment. This happens because the government has to accumulate assets to overcome its commitment problem and will therefore be richer in the long run. This allows labor taxes to be lower, which in turn increases labor supply. Higher labor will make capital more productive, which implies that capital will also be higher in steady state.

An interesting feature of the short run dynamics is that as long as commitment binds, capital may either be taxed or subsidized, depending on whether increasing capital makes the commitment constraint slacker or tighter. Numerical simulations will show an example where capital is being subsidized in the short run, so that the capital level is higher in the economy without commitment at all times.

On a more technical side, this paper provides a setup where the worst sustainable equilibrium can be determined in advance. Benhabib and Rustichini (1997) derive the best policy without commitment assuming that the worst punishment is known. Phelan and Stachetti (2001) argue that this is not always the case since the government's incentive constraint usually binds in the worst equilibrium, which means that the worst punishment has to be determined endogenously. This paper provides a sufficient condition for these two approaches to be equivalent. If the government is allowed to make lump sum transfers to consumers, which is a common assumption

in most taxation models, then it is always credible to give the households the worst possible expectations regarding future capital taxes, since it is incentive compatible for the government to tax the initial sunk capital at maximal rates, given that any remaining revenue can be redistributed to consumers as a lump sum transfer. Thus, no incentives need to be given for the government to act according to consumers' expectations, which means that the continuation of a worst equilibrium is still a worst equilibrium in this model, which allows us to determine the worst sustainable equilibrium in advance, as was assumed in Benhabib and Rustichini (1997).

We can interpret the long run results in this paper as an example of backloading of incentives, which is also present in models of commitment in other settings, such as Kocherlakota (1996), Ray (2002), or Acemoglu, Golosov and Tsyvinsky (2005). The idea is that in order to make the government's choice incentive compatible at all points in time, it is optimal to provide rewards as far off in the future as possible, since this provides incentives in all periods until then. Here, in particular, the backloading of incentives is achieved by letting the government increase its assets until the lack of commitment stops binding. This mechanism was not allowed by previous models that imposed budget balance.

This paper is also related to the work of Klein and Rios-Rull (2002), Klein, Krusell and Rios-Rull (2004), Klein, Quadrini and Rios-Rull (2005), and Klein, Krusell and Rios-Rull (2006), who look at time consistent Markov equilibria in taxation models. Since Markov equilibria preclude the use of trigger strategies, the set of equilibria that can be implemented is significantly smaller and in general a steady state with zero capital taxes will not be optimal even if the government is allowed to break budget balance. An exception to this is provided by Azzimonti-Renzo, Sarte and Soares (2006), where zero capital and labor income taxes are reached in the long run by collecting enough capital taxes in the initial periods to finance all future government spending. Although this paper reaches somewhat similar conclusions to ours, the mechanism at work is not the same. In Azzimonti-Renzo, Sarte and Soares (2006), capital accumulation occurs because in the short run capital is sunk, and it is in the government's best interest to use non distortionary taxation to finance future spend-

ing. As a consequence, asset accumulation will not stop until the government has enough assets to finance all future spending. Here, on the other hand, asset accumulation is used to make the future without default better, so that the incentive to default is reduced. Thus, asset accumulation stops when the incentive constraint for the government stops binding, which happens before the government's asset limit is reached, which means there is still positive labor taxation in the long run. Furthermore, along the transition path the predictions of the two models are significantly different, since here capital may even be subsidized in the short run if higher capital levels loosen the government's incentive constraint.

The game played between households and the government builds on the stream of literature developed by Chari and Kehoe (1990) on sustainable equilibria, which allows a more parsimonious definition of subgame perfect equilibria when some agents are too small to behave strategically. Chari and Kehoe (1993a and 1993b) use a setup without capital to model debt default. They allow for government default, but they either assume that households can commit to their debt, or debt repayment cannot be enforced at all. This paper, on the other hand, allows households to default, but it also allows the government to punish them if they do so, which makes household default non trivial.

The chapter proceeds as follows. Section 1.2 sets up the model with commitment and derives the optimal ex-ante plan for the government. Section 1.3 relaxes the assumption of commitment and characterizes the set of sustainable equilibria without commitment. Section 1.4 derives the best sustainable equilibrium under no commitment and analyses its long run properties. Section 1.5 presents a numerical example with short run dynamics and steady state results. Section 1.6 concludes with a brief summary of the main findings of the paper. Formal proofs are shown in the appendix in section 1.7.

1.2 Taxation with Commitment

This section introduces the economy with commitment. It characterizes allocations which are attainable under commitment for an arbitrary policy, which will also be relevant when there is no commitment since, from the households' perspective, they will be best responding to the government's strategy, which they take as given. The benchmark Chamley (1986) and Judd (1985) result is also derived.

1.2.1 Model Setup

The economy has a continuum of measure one of infinitely lived identical consumers, an arbitrary number of firms who behave competitively and a benevolent government. Time is discrete.

Households

The households' derive utility from consumption c_t , labor n_t , and consumption of a public good g_t . They discount the future at rate β , with $0 < \beta < 1$, so that each consumer's lifetime utility is given by

$$\sum_{t=0}^{\infty} \beta^t [u(c_t, n_t) + v(g_t)].$$

Assume u is increasing in consumption and decreasing in labor and globally concave. The usual Inada conditions hold $u_c(0, n) = \infty$, $u_c(\infty, n) = 0$, $u_n(c, 0) = 0$ and $u_n(c, \infty) = -\infty$. Assume also that the utility of the public good v is increasing and concave with $v'(0) = \infty$ and $v'(\infty) = 0$.

For each unit of work, households receive after tax wages of $w_t(1 - \tau_t^n)$. The labor tax can take any real value. Households can transfer consumption between periods using capital k_t or government bonds b_t . At time $t - 1$ households buy capital k_t . Each unit of capital costs one unit of consumption good. At time t households can rent this capital to firms for which they receive an after tax return of $R_t(1 - \tau_t^k)$. For simplicity, assume that capital is fully depreciated. If capital were depreciated at

rate $\delta < 1$, which may be irreversible, all the results in the paper remain unchanged. Steady state simulations will illustrate the effect of introducing irreversible capital. Assume households can always choose not to use their capital, so that capital taxes cannot be higher than one $\tau_t^k \leq 1$. No lower bound on τ_t^k is imposed. A bond that pays one unit of consumption good in period t costs q_{t-1} units of consumption good in period $t-1$. Households may also receive lump sum transfers from the government T_t , which must always be positive.

The households' per period budget constraint is thus given by

$$c_t + k_{t+1} + q_t b_{t+1} \leq R_t(1 - \tau_t^k)k_t + b_t + w_t(1 - \tau_t^n)n_t + T_t.$$

They must also meet the following no Ponzi condition

$$\lim_{t \rightarrow \infty} \left[b_{t+1} \prod_{s=0}^t q_s \right] \geq 0.$$

Government

The government is benevolent, which means that it maximizes the utility of a representative consumer. It needs to collect revenue to finance expenditure in the public good g_t every period. It sets proportional taxes on labor τ_t^n and capital τ_t^k each period. It transfers revenues between periods using government bonds b_t . The government sets the bond price q_t and consumers decide how many bonds to purchase. The government can make positive lump sum transfers to the households $T_t \geq 0$. Given this, the government's per period budget constraint is given by

$$g_t + b_t + T_t = w_t \tau_t^n n_t + R_t \tau_t^k k_t + q_t b_{t+1}.$$

Firms

Each period firms maximize profits given the before taxes prices for labor w_t and capital R_t . They have access to the production function $F(k_t, n_t)$, which has constant returns to scale and decreasing marginal productivity of capital and labor. Assume

$$F_k(0, n) = \infty, F_n(k, 0) = \infty, \text{ and } F_k(\infty, n) < 1/\beta.$$

Market equilibrium

Market must clear every period. For the goods market, this means that the resource constraint must be met every period

$$c_t + g_t + k_{t+1} = F(k_t, n_t).$$

Factor markets clear when factor prices equal the marginal productivity of each factor: $w_t = F_n(k_t, n_t)$ and $R_t = F_k(k_t, n_t)$.

1.2.2 Allocations Attainable under Commitment

Consider the commitment economy in which the government makes all its decisions for the future at the beginning of time. Households make their decisions after observing the policy plan decided by the government.

Let $\pi = (\pi_0, \pi_1, \dots)$ denote the sequence of government policies $\pi_t = (\tau_t^k, \tau_t^n, T_t, q_t, g_t)$, let $x = (x_0, x_1, \dots)$ denote the sequence of allocations $x_t = (c_t, n_t, k_{t+1}, b_{t+1})$, and let $p = (p_0, p_1, \dots)$ denote the sequence of market clearing prices $p_t = (R_t, w_t)$.

An allocation x is *attainable under commitment* if there are policies π and prices p such that (i) households maximize utility subject to their budget constraints and no Ponzi condition, (ii) the government meets its budget constraint with $T_t \geq 0$ and, (iii) factor prices equal marginal productivity of factors and the resource constraint is met.

Following the approach developed by Lucas and Stokey (1983), we can plug the first order condition for the households' problem into the households' budget constraint and obtain the economy's implementability condition. Lemma 1 shows that an allocation is attainable under equilibrium if and only if it meets the implementability condition and the resource constraint, as well as a transversality condition. Lemma 1 is proven in the appendix.

Lemma 1 *An allocation x is attainable under commitment if and only if it meets the following conditions for $t \geq 0$*

$$\begin{aligned} m(c_t, n_t) + \beta a_{t+1} &\geq a_t \\ c_t + g_t + k_{t+1} &= F(k_t, n_t) \\ \lim_{t \rightarrow \infty} \beta^t a_{t+1} &= 0 \end{aligned}$$

given k_0 and $a_0 = u_c(c_0, n_0)[F_k(k_0, n_0)(1 - \tau_0^k)k_0 + b_0]$, with $m(c_t, n_t)$ and a_t given by

$$\begin{aligned} m(c_t, n_t) &\equiv u_c(c_t, n_t)c_t + u_n(c_t, n_t)n_t \\ a_t &\equiv u_c(c_t, n_t) \left[\frac{u_c(c_{t-1}, n_{t-1})}{\beta u_c(c_t, n_t)} k_t + b_t \right] \text{ for } t > 0. \end{aligned}$$

Using a change of variables, we have replaced the level of government bonds for the value of consumer assets a_t . This will allow us to write the problem recursively using as state variables the level of capital and the value of consumer assets. This approach builds on Werning (2003), who rewrites the problem in Aiyagari, Marcet, Sargent and Seppala (2002) using the value of debt and the state of an exogenous Markov process as state variables.

The capital and labor income taxes associated with a given allocation x are determined by

$$\begin{aligned} \tau_{t+1}^k &= 1 - \frac{1}{\beta F_k(k_{t+1}, n_{t+1})} \frac{u_c(c_t, n_t)}{u_c(c_{t+1}, n_{t+1})} \\ \tau_t^n &= 1 + \frac{1}{F_n(k_t, n_t)} \frac{u_n(c_t, n_t)}{u_c(c_t, n_t)}. \end{aligned}$$

We can guarantee that the transversality condition $\lim \beta^t a_{t+1} = 0$ is met by constraining a to always be below the natural debt limit $\bar{a}(k_t)$ which is the maximum debt level that can be repaid by the government

$$\bar{a}(k_t) \equiv \max_{c, n, k} \sum_{s=t}^{\infty} \beta^{s-t} m(c_s, n_s) \text{ st } c_s + k_{s+1} \leq F(k_s, n_s).$$

From now on the implementability condition and the resource constraint will be used as necessary and sufficient conditions for an allocation to be attainable under commitment with the underlying condition that a_t must remain below this upper bound.

This characterization of allocations attainable under commitment will be useful to determine the optimal policies with or without commitment since in both cases the resulting allocations will have to be chosen by households who anticipate a given set of policies, which means that their outcomes must be attainable under commitment.

1.2.3 Optimal Taxes with Commitment

This section derives the optimal policy plan when the government can choose the policies for all future periods at time zero. It introduces a recursive formulation of the problem (that will also be used for the no commitment case) to derive the benchmark Chamley (1986) and Judd (1985) result of zero capital taxes in the long run.

The Ramsey problem chooses among all the allocations attainable under commitment, the one that maximizes the welfare of the representative consumer. The outcome of a *Ramsey equilibrium* is a sequence x that maximizes the present value of utility $\sum_{t=0}^{\infty} \beta^t [u(c_t, n_t) + v(g_t)]$ subject to x being attainable under commitment and given an initial stock of capital $k_0 > 0$ and an initial promise for the value of consumer assets $a_0 < \bar{a}(k_0)$. This formulation assumes that the initial planner has committed to a given a_0 . If instead we wanted to assume that the government had an initial outstanding debt of b_0 we would have to add the following restriction for the initial period $a_0 = u_c(c_0, n_0)[F_k(k_0, n_0)(1 - \tau_0^k)k_0 + b_0]$. Given this, we can write the Ramsey problem using the following sequence formulation

$$\begin{aligned}
 V(k_0, a_0) &\equiv \max_{c, n, g, k, a} \sum_{t=0}^{\infty} \beta^t [u(c_t, n_t) + v(g_t)] \\
 \text{subject to} & \quad m(c_t, n_t) + \beta a_{t+1} \geq a_t \\
 & \quad c_t + g_t + k_{t+1} = F(k_t, n_t).
 \end{aligned}$$

Using a and k as state variables this problem can be written recursively in the following way

$$\begin{aligned} V(k, a) &= \max_{c, n, g, k, a} [u(c, n) + v(g) + \beta V(k', a')] \\ \text{subject to} \quad & m(c, n) + \beta a' \geq a \\ & c + g + k' \leq F(k, n). \end{aligned}$$

If $m(c, n)$ is concave, then the constraint set is convex, which means that the value function $V(k, a)$ will be concave, and the first order conditions are necessary and sufficient for optimality. If this condition is not met, first order conditions are still necessary for an optimum, but no longer sufficient, since it is also necessary to verify that the second order conditions are met to make sure we are at a maximum.

Using μ as the multiplier on the implementability condition and ρ as the multiplier on the resource constraint, we can write the Lagrangean for this problem in the following way

$$L = u(c, n) + v(g) + \beta V(k', a') + \mu[m(c, n) + \beta a' - a] - \rho[c + g + k' - F(k, n)].$$

Combining the first order conditions for k' and a' with the envelope conditions for k and a , we get the following equations

$$\begin{aligned} V_k(k, a) &= \beta F_k(k, n) V_k(k', a') \\ V_a(k', a') &= V_a(k, a). \end{aligned}$$

A steady state for the Ramsey economy has constant c , n , g , k and a as well as constant multipliers μ and ρ . From the optimality condition for k , it is clear that in steady state $\beta F_k = 1$. Plugging into the expression for capital taxes, it is straightforward to see that capital taxes must be zero in steady state

$$\tau^k = 1 - \frac{1}{\beta F_k} \frac{u_c}{u'_c} = 0.$$

Labor taxes, on the other hand, will remain positive in steady state

$$\tau_t^n = 1 + \frac{1}{F_n} \frac{u_n}{u_c}.$$

This is the well known Chamley (1986) and Judd (1985) result that capital income taxes converge to zero in the long run.

1.3 Sustainable Equilibria without Commitment

This section introduces lack of commitment by modelling the taxation problem as a game where the government and the agents in the economy make sequential decisions every period. The equilibrium concept is defined and a simple characterization of equilibrium outcomes is formalized based on a maximum threat point of reversion to the worst equilibrium, in the spirit of Abreu (1988).

1.3.1 Game Setup

This section introduces a game where the lack of commitment is modelled explicitly and the value of default that sustains the initial plan is determined endogenously. Since households and firms behave competitively, whereas the government behaves strategically, I will use the notion of sustainable equilibria introduced by Chari and Kehoe (1990), where all strategies are conditional on the past history of the government's actions.

Chari and Kehoe (1993a and 1993b) model debt default in a setup without capital. Both papers allow the government to default on its bonds. In Chari and Kehoe (1993a) it is assumed that households can always commit to repay their debt. Conversely, in Chari and Kehoe (1993b) it is assumed that households cannot commit to their debt, which means that debt repayment cannot be enforced at all, leading to no loans being made to households in equilibrium. By introducing an endogenous punishment for default, we now make the households' default decision non trivial since they will only default if they expect not to get punished harshly enough.

Assume that the government cannot commit to future taxes and transfers. Furthermore, both households and the government can default on their bonds. The government can punish consumers who defaulted. Namely, each period the government chooses $P_t \geq 0$, which is the utility loss that consumers who defaulted in the previous period experience. Let d_t^g be an indicator function for whether the government defaults and d_t^c be the percentage of consumers who defaulted in period t .

The timing of the game is as follows. At the beginning of the period, the government decides by how much to punish consumers who defaulted in the previous period, whether to default on its bonds, and which taxes and transfers to set for the current period. Note that the government does not need to know which households defaulted; it is sufficient that the government knows the percentage of households who defaulted. Since each household has mass zero, the action of a finite number of households does not affect the percentage of households defaulting, which means that they are still non strategic. To put the punishment into practice the government will then rely on an outside independent institution (maybe courts) which will be able to punish each defaulting household.

After the government has made its choices, allocations and prices are jointly determined by the households' and the firms' maximization problems at market clearing prices.

1.3.2 Strategies

The government's actions in period t now include the punishment and decision to default, so that the expanded vector of government's actions is now $\Pi_t = (\tau_t^k, \tau_t^n, T_t, q_t, g_t, d_t^g, P_t)$.

Each period, every household chooses how much to consume, work and invest in capital and bonds. It also decides whether to default on its debt. Let d_t^{ci} be an indicator function for whether household i chooses to default in period t . The vector of individual decisions in each period is $X_t^i = (c_t^i, n_t^i, k_{t+1}^i, b_{t+1}^i, d_t^{ci})$. The vector of aggregate choices that results from the households decisions is $X_t = (c_t, n_t, k_{t+1}, b_{t+1}, d_t^c)$, where the aggregate value of each aggregate variable is the integral over all the house-

holds in the economy of the individual variables. In equilibrium, since all households are identical and follow pure strategies, the aggregate action will be the same as each individual action.

The price vector is $p_t = (R_t, w_t)$ as before.

Let h_t be the history of government decisions until time t so that $h_t = (\Pi_0, \dots, \Pi_t)$. Following Chari and Kehoe (1990), all the strategies in the game will be contingent only on this history, since households are infinitesimal and have no power to influence X_t , which means that they will not behave strategically. Thus, knowing the households' strategies and government's actions until time t is enough to characterize all the history until then.

The strategy for the government is given by σ . The strategy for each period t is a mapping from the history h_{t-1} into the government's decision space Π_t , so that $\Pi_t = \sigma_t(h_{t-1})$. When choosing a given strategy, the government anticipates that histories will evolve according to $h_t = (h_{t-1}, \sigma_t(h_{t-1}))$. Let σ^t denote the sequence of government strategies from time t onwards.

The strategy for a representative household is given by f . The strategy for each period t is a mapping from the history h_t into the households' decision space X_t , so that $X_t = X_t^i = f_t(h_t)$. Let f^t denote the sequence of household strategies from time t onwards.

Firms and markets jointly work as a third player that has strategy ϕ mapping the history h_t into the vector of factor prices p_t , so that $p_t = \phi(h_t)$. Let f^t denote the sequence of household strategies from time t onwards.

In the next section we will specify how each player chooses its strategy in a sustainable equilibrium.

1.3.3 Sustainable Equilibria

At time t , the government and households choose an action for time t and a contingent plan for the future. This is equivalent to choosing an action for today while anticipating future behavior since both the government and households have time consistent preferences, which means that the plan they choose today will be optimal tomorrow.

The problem solved by the government and households at time t is described below.

For every history h_{t-1} , given allocation rule f and pricing rule ϕ , the government chooses σ^t to maximize the present value of utility

$$\sum_{s=t}^{\infty} \beta^{s-t} [u(c_s(h_s), n_s(h_s)) + v(g_s(h_{s-1})) - d_t^c(h_{s-1})P_s(h_{s-1})]$$

subject to

$$\begin{aligned} g(h_{s-1}) + T_s(h_{s-1}) &= w_s(h_s)\tau_s^n(h_{s-1})n_s(h_s) + R_s(h_s)\tau_s^k(h_{s-1})k_s(h_{s-1}) + \\ &\quad q_t(h_{s-1})b_{s+1}(h_s) - b_s(h_{s-1})(1 - d_t^c(h_s))(1 - d_t^g(h_{s-1})) \\ T_s(h_{s-1}) &\geq 0 \end{aligned}$$

and realizing that future histories are induced by σ^t according to $h_s = (h_{s-1}, \sigma_s(h_{s-1}))$.

For every history h_t , given policy rule σ (and the histories it induces) and pricing rule ϕ , each household chooses f^t to maximize the present value of utility

$$\sum_{s=t}^{\infty} \beta^{s-t} [u(c_s^i(h_s), n_s^i(h_s)) - d_t^{ci}(h_{s-1})P_s(h_{s-1})]$$

subject to

$$\begin{aligned} c_s^i(h_s) &= w_s(h_s)(1 - \tau_s^n(h_{s-1}))n_s^i(h_s) + T_s(h_{s-1}) + \\ &\quad b_s^i(h_{s-1})(1 - d_t^{ci}(h_s))(1 - d_s^g(h_{s-1})) - q_s(h_{s-1})b_{s+1}^i(h_s) + \\ &\quad R_s(h_s)(1 - \tau_s^k(h_{s-1}))k_s^i(h_{s-1}) - k_{s+1}^i(h_s). \end{aligned}$$

Market clearing and firm optimality require that for every history h_t firm demand must equal household supply for every production factor, which happens when factor prices equal their marginal productivity, so that $\phi_t(h_t)$ is given by

$$\begin{aligned} w_t(h_t) &= F_n(k_t(h_{t-1}), n_t(h_t)) \\ R_t(h_t) &= F_k(k_t(h_{t-1}), n_t(h_t)). \end{aligned}$$

A *sustainable equilibrium* is a triplet (σ, f, ϕ) that satisfies the following conditions: (i) given f and ϕ , the continuation of contingent policy plan σ solves the government's problem for every history h_{t-1} ; (ii) given σ and ϕ , the continuation of contingent allocation rule f solves the households' problem for every history h_t ; (iii) given f and σ , the continuation of the contingent pricing rule ϕ is such that factor prices equal marginal productivity for every history h_t .

1.3.4 Worst Sustainable Equilibrium

Let $V(\sigma, f, \phi)$ denote the present value of utility that results from a sustainable equilibrium (σ, f, ϕ) . Then the worst sustainable equilibrium is the sustainable equilibrium that leads to the lowest value $V(\sigma, f, \phi)$. It will be useful to find the worst sustainable equilibrium to then define which equilibria can be sustained without ex-ante commitment, since the worst sustainable equilibrium is the worst punishment that can be credibly inflicted on a government that deviates from a predefined plan.

Lemma 2 *The value of the worst sustainable equilibrium only depends on the current capital level: $V(\sigma^w, f^w, \phi^w) = V^w(k)$.*

A proof of this lemma can be found in the appendix. The idea is that the value of a worst equilibrium can only depend on the current payoff relevant variables. Furthermore, since the government can eliminate debt by defaulting, no additional punishment can be given to it, which means that the value of the worst sustainable equilibrium will not depend on the level of debt.

For our model the worst sustainable equilibrium is an equilibrium where all agents default on their debt and the government always expropriates capital.

The *default equilibrium* is a triplet (σ^d, f^d, ϕ^d) where agents have the following strategies:

(i) The government always defaults on b_t , never punishes consumers, always taxes capital at confiscatory rates, and sets $q_t = 0$. Transfers and labor taxes implement

the solution to the following problem

$$\begin{aligned} & \max_{c_t, n_t, g_t} [u(c_t, n_t) + v(g_t)] \\ & \text{subject to} \quad m(c_t, n_t) \geq 0 \\ & \quad \quad \quad c_t + g_t = F(k_t, n_t). \end{aligned}$$

(ii) Households default if $b_t < 0$, never invest in capital, and never lend or borrow from the government. Labor and consumption solve the following problem

$$\begin{aligned} & \max_{c_t, n_t} u(c_t, n_t) \\ & \text{subject to} \quad c_t = T_t + w_t(1 - \tau_t^n)n_t. \end{aligned}$$

(iii) Factor prices equal marginal productivity of factors.

Lemma 3 *The default equilibrium is the worst sustainable equilibrium:*

$$V(\sigma^w, f^w, \phi^w) = V^w(k) = V^d(k).$$

The worst sustainable equilibrium punishes the government by giving households beliefs about the government's future behavior that lead to low future value. However, these beliefs have to be correct, so the government has to be given incentives to keep the plan. Phelan and Stachetti show that in general the government's incentive constraint will be binding, which means the continuation value of a worst sustainable equilibrium will not be a worst sustainable equilibrium. However, since we allow the government to make lump sum transfers to the households, even when the government gives households extremely pessimistic beliefs that capital will be fully expropriated in the following period, the government's incentive constraint will not bind since it is always willing to tax capital at maximal rates and then redistribute back to the households.

Since households always expect the government to fully expropriate capital, they will never invest even though capital is very productive. Furthermore, if any household actually invests, it will be in the government's best interest to expropriate it

since capital is sunk ex-post. Thus, this lack of commitment will lead to an extremely inefficient investment decision. Since both households and the government are best responding to each other's strategy, the default equilibrium is sustainable. The appendix proves that the default equilibrium is the worst sustainable equilibrium.

Since the best the government can do when households are playing a default equilibrium is to maximize its per period utility, the flow utility reached in a period when the initial stock of capital is k is given by

$$U^d(k) = \max_{c,n,g} [u(c,n) + v(g)] \text{ st } m(c,n) \geq 0, c + g = F(k,n)$$

and the net present value of a default equilibrium is

$$V^d(k) = U^d(k) + \frac{\beta}{1-\beta} U^d(0),$$

which only depends on the initial level of k .

1.3.5 Characterization of Sustainable Outcomes

In the spirit of Abreu's optimal punishments, we will use reversion to the worst sustainable equilibrium as the maximum threat point that allows us to sustain equilibria. Thus, for an equilibrium to be sustainable, it must yield higher utility than the worst sustainable equilibrium in all future dates. We are using the terminology of Abreu (1988), who defines optimal punishments when firms deviate in a cartel. Here, there is not any kind of collusion *per se*, but we still need to enforce cooperation since the government's ex-ante and ex-post incentives are not aligned. Thus, the worst punishment is not inflicted by other firms, but rather by changing the consumer's expectations, which leads to a different equilibrium that is worse for everyone.

The next lemma characterizes the entire set of sustainable equilibrium outcomes, which are the allocations that are induced by a particular sustainable equilibrium.

Lemma 4 *An allocation x is the outcome of a sustainable equilibrium if and only if:*

- (i) *x is attainable under commitment.*

(ii) *The continuation value of x is always better than the worst sustainable equilibrium*

$$\sum_{t=i}^{\infty} \beta^{t-i} [u(c_t, n_t) + v(g_t)] \geq V^d(k_i).$$

The proof of this lemma can be found in the appendix. The idea is that for an allocation to be the outcome of a sustainable equilibrium, it is necessary that households and firms are optimizing given the government's strategy, which means that the resulting allocation must be attainable under commitment. Government optimality requires that it is never in the government's best interest to deviate. Since the worst punishment after a deviation is $V^d(k)$, this gives us a lower bound on the utility that can be reached in a sustainable equilibrium at any point in time.

1.4 Best Sustainable Equilibrium

Now that the set of sustainable equilibria has been characterized, we turn to finding among these, the one that maximizes the initial welfare for the society for given initial conditions for k_0 and a_0 . As before, assuming commitment to a given a_0 in the initial period is a simplifying assumption that can be relaxed. Assume that a_0 is within the necessary bounds for an equilibrium to exist.

1.4.1 Sequence Approach

The outcome of the best sustainable equilibrium solves

$$\begin{aligned} V(k_0, a_0) &= \max_{c, n, g, k, a} \sum_{t=0}^{\infty} \beta^t [u(c_t, n_t) + v(g_t)] \\ \text{subject to} \quad & m(c_t, n_t) + \beta a_{t+1} \geq a_t \\ & c_t + g_t + k_{t+1} = F(k_t, n_t) \\ & \sum_{t=i}^{\infty} \beta^{t-i} [u(c_t, n_t) + v(g_t)] \geq V^d(k_i). \end{aligned}$$

This follows directly from lemma 4. Since the three restrictions are necessary and sufficient for a sustainable equilibrium, then the allocation that maximizes welfare

subject to them must be the outcome of the best sustainable equilibrium.

This formulation is equivalent to the Ramsey problem under commitment, with an additional incentive compatibility condition that ensures that, at each point in time, the government never wants to deviate from the predefined plan.

Notice that one of our restrictions now has an endogenous function $V^d(k_i)$, which is concave. Given this, the constraint set may not be convex, even if we assumed that $m(c_t, n_t)$ is concave. Thus, we cannot guarantee that $V(k, a)$ is a concave function. In the analysis that follows we proceed as if $V(k, a)$ were concave. The appendix shows that the same results follow through even if that is not the case.

Let $x^{fb}(k_0) \equiv \arg \max \sum_{t=0}^{\infty} \beta^t [u(c_t, n_t) + v(g_t)]$ st $c_t + g_t + k_{t+1} = F(k_t, n_t)$ be the first best allocation when the economy has initial capital k_0 .

Then the implementability constraint will not be binding for $a_0 \leq \underline{a}(k_0)$, which is defined by $\underline{a}(k_0) \equiv \sum_{t=0}^{\infty} \beta^t m(c_t^{fb}(k_0), n_t^{fb}(k_0))$.

Notice that $V^{fb}(k) \equiv V(k, \underline{a}(k)) > V^d(k)$, which means that if the implementability condition is not binding, then the incentive compatibility condition will not bind either. The reason for this is straightforward: if the government does not have to use distortionary taxation to finance its spending, then it has no incentive to deviate from the optimal plan.

Since the implementability condition is not binding for $a_0 \leq \underline{a}(k_0)$, then having lower a_0 does not bring any additional benefit to the economy, which means that $V_a(k, a) = 0$ for $a \leq \underline{a}(k)$. On the other hand, if $a > \underline{a}(k)$, then starting off with a lower a_0 relaxes the implementability condition, which implies that $V_a(k, a) < 0$ for $a > \underline{a}(k)$.

1.4.2 Recursive Approach

The program to find the best sustainable equilibrium can also be written recursively as stated below. The appendix shows that the two formulations are equivalent, and from now on the recursive approach will be used.

$$\begin{aligned}
V(k, a) &= \max_{c, n, g, k, a} [u(c, n) + v(g) + \beta V(k', a')] \\
\text{subject to} & \quad m(c, n) + \beta a' \geq a \\
& \quad c + g + k' = F(k, n) \\
& \quad V(k', a') \geq V^d(k').
\end{aligned}$$

The Lagrangean for this problem is

$$\begin{aligned}
L &= u(c, n) + v(g) + \beta V(k', a') + \mu[m(c, n) + \beta a' - a] \\
& \quad - \rho[c + g + k' - F(k, n)] + \gamma\beta[V(k', a') - V^d(k')].
\end{aligned}$$

Combining the first order conditions for k' and a' with the envelope conditions for k and a , we get the following equations

$$\begin{aligned}
V_k(k', a') &= \frac{V_k(k, a)}{\beta F_k(k, n)} + \gamma[V_k^d(k') - V_k(k', a')] \\
V_a(k', a')(1 + \gamma) &= V_a(k, a).
\end{aligned}$$

The optimality condition for k shows how the lack of commitment can distort the choice of capital in the short run. If commitment binds ($\gamma > 0$) and the value of default reacts more to changes in capital than the value of the optimal sustainable plan, then capital will be distorted downward, since this will help loosen the incentive compatibility constraint. If conversely, the value of the optimal sustainable plan varies more with capital, than capital will be distorted upward. Thus, if higher capital makes commitment less binding, it will be optimal to subsidize capital. This result is reminiscent of the findings of Benhabib and Rustichini (1997), who show that it could be optimal to either tax or subsidize capital. However, here this will only be true in the short run, since in the long run the economy will converge to a steady state where commitment does not bind.

The optimality condition for a says that it is optimal for the value of government

assets to decrease over time as long as commitment is binding, which leads to an increase of government assets over time. The reason for this is that when the government accumulates assets, it gets a direct benefit of higher utility tomorrow, as well as an additional benefit from loosening the incentive compatibility condition in the future. Thus, to some extent, government assets work as a commitment mechanism that reduces the incentive to default by increasing the welfare of the equilibrium strategy.

The next section describes the long run properties of the economy without commitment.

1.4.3 Steady State

In this section we derive the paper's main result that capital taxes must converge to zero in the long run. We start by showing that any steady state must have zero capital taxes and then show that the economy will indeed converge to a steady state.

Proposition 1 (*Zero capital taxes in steady state*) *In steady state, the best sustainable equilibrium has zero capital taxes.*

Proof. Assume the economy is in a steady state with constant c , n , g , k , and a . The first order conditions for c and n ($u_c + \mu m_c = \rho$ and $u_n + \mu m_n = -\rho/F_n$) imply that ρ and μ must also be constant. We can now prove by contradiction that capital taxes cannot be different from zero in the long run.

If capital taxes are not zero, then $\beta F_k(k, n) \neq 1$. From the optimality condition for k derived above, this implies that $\gamma > 0$ (recall that $\gamma \geq 0$ since it is the multiplier on an inequality constraint). We can see in the optimality condition for a that when $\gamma > 0$, it must be true that $V_a(k, a) = 0$ is equal to zero in steady state. But then it must be true that $a \leq \underline{a}(k)$ and $V(a, k) = V^{fb}(k) > V^d(k)$, which means that the incentive compatibility condition cannot be binding and we must have $\gamma = 0$, which means we have reached a contradiction and capital taxes cannot be different from zero in steady state. ■

Proposition 2 (Positive labor taxes in steady state) *In steady state, the best sustainable equilibrium has positive labor taxes as long as $a_0 > \underline{a}(k_0)$.*

Proof. Let $\bar{a}(k)$ be the highest value of a such that the incentive compatibility condition is met, which is defined by $V(\bar{a}(k), k) = V^d(k)$. Since $V^{fb}(k) = V(\underline{a}(k), k)$ and $V^{fb}(k) > V^d(k)$, it must be the case that the following inequality holds

$$V^{fb}(k) = V(\underline{a}(k), k) > V(\bar{a}(k), k) = V^d(k).$$

Since $V_a = 0$ for $a < \underline{a}(k)$ and $V_a < 0$ for $a > \underline{a}(k)$, this implies that $\bar{a}(k) > \underline{a}(k)$, which means that the incentive compatibility condition stops binding before natural asset limit is achieved.

In steady state, the incentive compatibility condition is either exactly met, or it is slack. If the incentive compatibility is exactly met, then $a = \bar{a}(k) > \underline{a}(k)$, which means that $\mu > 0$ since $\underline{a}(k)$ is the highest value of a for which the implementability condition does not bind. If incentive compatibility is slack, then given that $\mu' = (1 + \gamma)\mu$, it must be the case that μ is constant since the last time the incentive compatibility condition binded (or since the initial period, if it never binded), where it was strictly positive by the argument above. ■

Proposition 3 (Convergence to steady state) *The best sustainable equilibrium converges to a steady state.*

Proof. The first order conditions describe the unique path for the economy for given initial conditions.

If γ converges to zero, then the long run dynamics of capital and a are like those of an economy without commitment, which are governed by

$$\begin{aligned} V_k(k', a') &= \frac{V_k(k, a)}{\beta F_k(k, n)} \\ V_a(k', a') &= V_a(k, a). \end{aligned}$$

This means that capital will increase as long as it is below its steady state level ($\beta F_k > 1$) and decrease when it is above the steady state level ($\beta F_k < 1$), so that it

converges to its steady state.

If γ does not converge to zero, then $V_a(k, a)$ must converge to zero since γ is weakly positive and $V_a(k', a')(1 + \gamma) = V_a(k, a)$. But we have just seen that when $V_a(k, a)$ is close to zero, the implementability condition cannot be binding, which means that γ must converge to zero. ■

As long as commitment is binding, increasing government savings not only increases tomorrow's continuation value, but also loosens the incentive compatibility constraint. Thus, the government will keep saving until it has achieved enough assets for the incentive compatibility to stop binding. This will happen before the government reaches its asset limit, since $V(\underline{a}(k), k) > V^d(k)$.

We have seen so far that without commitment the economy will still converge to a steady state where commitment does not bind. Next we explore the long run implications that the lack of commitment may have.

With commitment, a steady state had to meet the following conditions

$$\begin{aligned} u_n + \mu m_n &= -\rho F_n & u_c + \mu m_c &= \rho \\ c + g + k &= F(k, n) & v_g &= \rho \\ \mu[m(c, n) - a(1 - \beta)] &= 0 & \beta F_k &= 1. \end{aligned}$$

Without commitment, any steady state commitment still has to meet the previous conditions, but it also has to meet the incentive compatibility condition

$$V(k, a) = \frac{1}{1 - \beta}[u(c, n) + v(g)] \geq V^d(k).$$

Thus, all steady states without commitment are also steady states under commitment. However, the converse need not be true. In particular, steady states with a very indebted government (which translates into a high level of a) need to use more distortionary taxation, which reduces the present value of utility, so that the incentive compatibility condition may not hold. As one would expect, more indebted

governments have a higher incentive to default. Conversely, steady states where the government has a substantial amount of assets do not need to use much distortionary taxation, which reduces the incentive to default and expropriate capital.

1.5 Numerical Simulations

To illustrate the results of the model, consider an economy with preferences given by

$$u(c, n) + v(g) = \ln(c) - n^{1+\theta}/(1 + \theta) + \ln(g)$$

and production function

$$F(k, n) = Ak^\alpha n^{1-\alpha} + Bn + Ck,$$

where the specific parameters are $\theta = 1$, $\beta = 0.4$, $A = 2$, $B = 0.5$, $C = 0.5$, and $\alpha = 0.5$.

We will start by looking at the short run dynamics of this economy for given initial conditions, in order to see how the lack of commitment changes the path towards the steady state and what welfare costs it entails.

Next, we turn too see how commitment changes the feasible steady states of this economy, in order to infer what the long run implications of lack of commitment are.

1.5.1 Short Run Dynamics

The recursive problem can be written as

$$\begin{aligned} V(k, a) &= \max_{k', a'} \{U(k, a, k' a') + \beta V(k', a')\} \\ \text{subject to} & \quad V(k', a') \geq V^d(k') \end{aligned}$$

where $U(k, a, k'a')$ is given by

$$\begin{aligned}
 U(k, a, k'a') &= \max_{c, n, g} [u(c, n) + v(g)] \\
 \text{subject to} & \quad m(c, n) + \beta a' \geq a \\
 & \quad c + g + k' = F(k, n).
 \end{aligned}$$

The first step of the simulation is to construct $U(k, a, k'a')$. For computational convenience, it is useful to stack the state variables for each period into one single dimension. If we exclude initial conditions where the implementability condition is not binding, we can assume, without loss of generality, that the implementability condition holds with equality. Given this and the previous assumptions for the functional form, we can write the implementability condition as $1 - n^{1+\theta} + \beta a' \geq a$, which means we can solve for $n(k, a, k'a')$, which is a matrix that gives us the optimal choice of n for a given state today on one axis and a given state tomorrow on the other axis. Given the functional form for utility it is optimal to have $g = c$. Thus, we can use the resource constraint to find $g(k, a, k'a')$ and $c(k, a, k'a')$. Finally, we can use these matrices to compute $U(k, a, k'a') = u[c(k, a, k'a'), n(k, a, k'a')] + v[g(k, a, k'a')]$.

Given this, we can make an initial guess for $V(k, a)$ and iterate on it until we find a fixed point. For the case where there is no commitment, this iteration includes a large punishment when the incentive compatibility constraint is violated. This ensures that for the chosen solution the constraint is always met.

Figure 1-1 shows the optimal path for an economy with initial capital stock $k_0 = 0.1$ and an initial promise for the value of consumer assets $a_0 = 1.3$. Dashed lines represent the economy with commitment, whereas the solid lines represent the no-commitment economy.

For the parameter values we have assumed, capital is being subsidized when the incentive compatibility is binding. This reflects the fact that the value of the best equilibrium is more sensitive to changes in the level of capital than the worst equilibrium. Thus, increasing capital relaxes the incentive compatibility condition.

In the economy without commitment, the value of a also decreases in initial periods

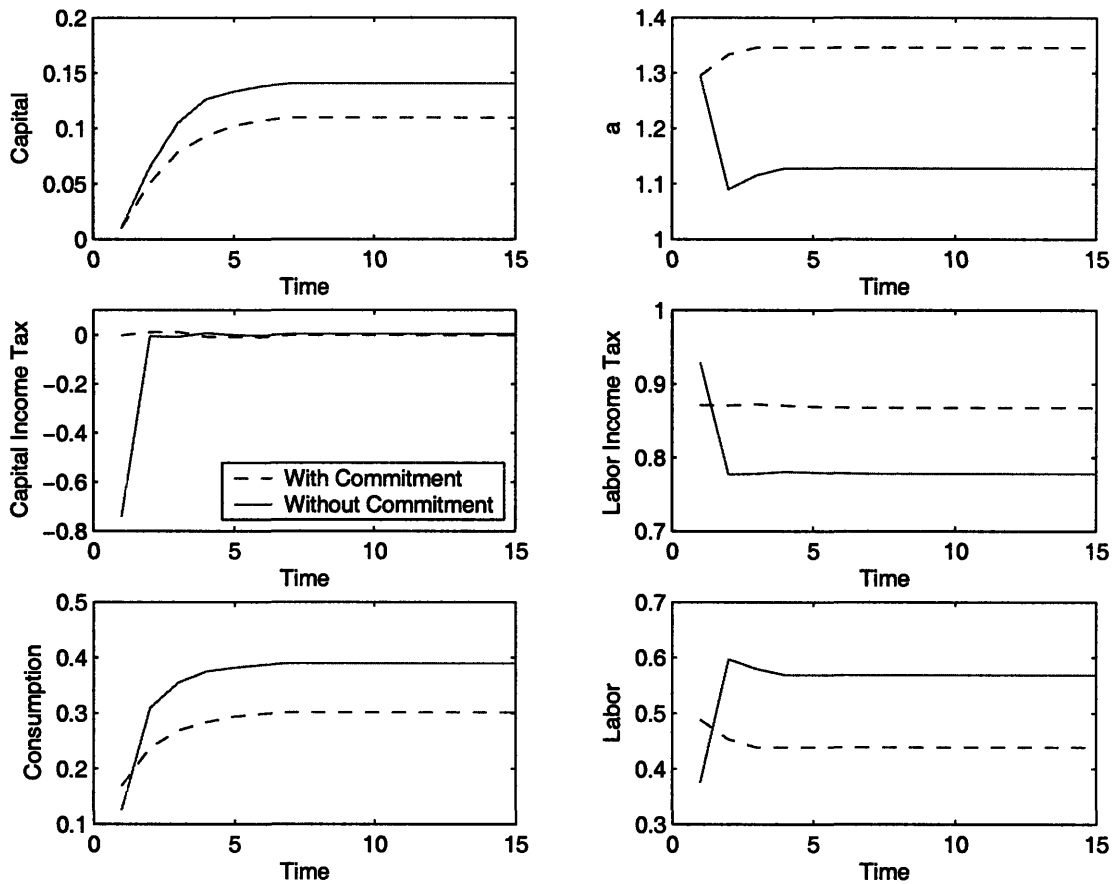


Figure 1-1: Transition Dynamics with and without Commitment

(which is equivalent to reducing the government's debt) so that, in the future, labor taxes are lower and labor and capital levels are higher without commitment.

As a consequence, capital levels are higher in the economy without commitment both in the short run (due to the capital subsidy) and in the long run (due to lower labor taxes).

The initial value of welfare in the economy with commitment is -5.57 , whereas, for the economy without commitment, initial welfare is -5.82 . This difference represents the cost of lack of commitment.

In the long run, however, the welfare in the economy without commitment (-3.41) is higher than in the economy with commitment (-4.08), since the economy without

commitment reduced its debt in the initial periods, which means that it does not have to do as much distortionary taxation in steady state. Thus, while there is a short run cost to commitment, in the long run, an economy whose government cannot commit ex-ante may actually have higher welfare.

1.5.2 Long Run Implications

In the long run, commitment will not be binding even if the government cannot commit to policies ex ante. Section 4.3 shows that steady states without commitment are also steady states in the economy with commitment, albeit with different initial conditions. However, the converse need not be true since, without commitment, we cannot sustain steady state equilibria of the economy with commitment where the incentive compatibility condition is not met. Thus, introducing the inability of governments to commit can help us predict what kinds of equilibria we might expect to find.

Figure 1-2 plots the set of steady state equilibria under commitment, indexed by the steady state level of capital. Notice that steady states with higher capital also have lower a , which means that the government is less indebted. Thus, $V(k, a)$ increases with capital not only because capital is increasing but also because a is decreasing. Labor and consumption are increasing with capital, exactly because steady states with higher capital have lower labor taxes.

To find which steady states are still feasible when the government cannot commit to future policies, the top panels plot the value of default associated with the capital level that is chosen in each of the equilibria. The left panel shows the value of default in our baseline model, whereas the right panel plots the value of default when the depreciated capital cannot be reconverted into consumption goods. As would be expected, in the economy with irreversible capital, the value of default is higher, but the qualitative results do not change¹. Without commitment, steady states must meet

¹If capital depreciates at rate δ and the non-depreciated capital is irreversible then capital income taxes still converge to zero in the long run in the best sustainable equilibrium. The more substantial difference is that default is now more attractive, since in the default equilibrium the level capital remains positive and only decreases as it depreciates, which does not allow such as stark punishment

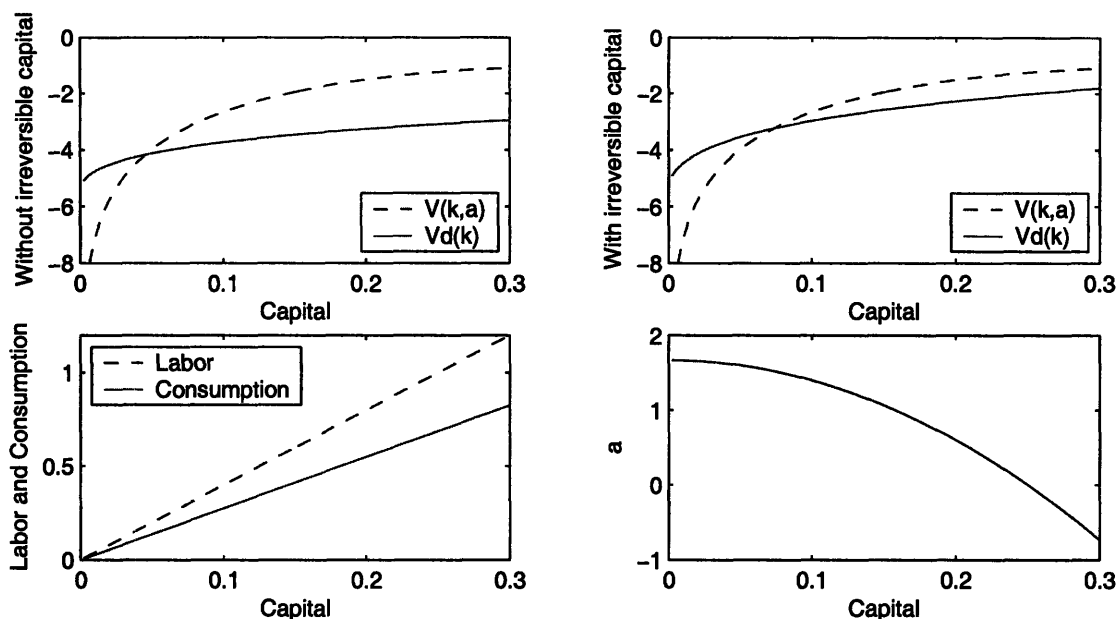


Figure 1-2: Steady States achievable with and without Commitment

the incentive compatibility constraint $V(k, a) \geq V^d(k)$, which means that only steady states with a high enough level of capital are sustainable. In general, steady states where the government is very indebted will not be feasible without commitment. In particular, when governments approach their natural debt limit and need to tax labor income at very distortionary levels, it becomes more likely that they will default on their past promises.

as before. Let the value of a default equilibrium in this case be given by $V^D(k)$.

The outcome of the best sustainable equilibrium solves

$$\begin{aligned}
 V(k, a) &= \max_{c, n, g, k, a} [u(c, n) + v(g) + \beta V(k', a')] \\
 \text{subject to} \quad & m(c, n) + \beta a' \geq a \\
 & c + g + k' = F(k, n) + (1 - \delta)k_t \\
 & k_{t+1} \geq (1 - \delta)k_t \\
 & V(k', a') \geq V^D(k')
 \end{aligned}$$

Since $V^D(k)$ is still lower than $V^{fb}(k)$ capital taxes are still zero in the long run.

The value for depreciation is $\delta = 1 - C$, so that the steady state equilibria with commitment remain unchanged with or without irreversibility of capital.

1.6 Concluding Remarks

If the government is allowed to accumulate assets, then, even if it cannot commit to a stream of future taxes, in the long run capital taxes will converge to zero. The previous literature on taxation without commitment had found that this was not the case when the government was forced to keep budget balance in every period. The reason for this disparity is that, in the absence of commitment, government assets can work as a commitment device to discipline government behavior that was not available under budget balance.

Thus, in the short run, economies without commitment will tend to accumulate more assets. Furthermore, while commitment binds, capital may either be taxed or subsidized, depending on whether increasing capital loosens or tightens the government's incentive constraint.

In the long run, economies without commitment will tend to have a higher asset level, which leads to higher capital levels since an economy with a richer government will have lower labor taxes. This in turn increases labor, leading to higher productivity of capital.

1.7 Appendix

1.7.1 Proof of Lemma 1

The intertemporal budget constraint can be written as

$$R_0(1 - \tau_0^k)k_0 + b_0 \leq \sum_{t=0}^{\infty} \left[(c_t - w_t(1 - \tau_t^n)n_t - T_t) \prod_0^t q_s \right] \\ + \lim_{t \rightarrow \infty} \left[\left(\frac{k_{t+1}}{q_t} + b_{t+1} \right) \prod_0^t q_s \right].$$

Since the marginal utility of consumption is always positive, it will always be optimal for households to meet their budget constraint with equality and to choose the long run values of k and b such that the last term in the budget constraint is not

positive, which would make it more binding.

Together with the no Ponzi condition and the non negativity constraint on capital, this implies that the following transversality condition must be met

$$\lim_{t \rightarrow \infty} \left[\left(\frac{k_{t+1}}{q_t} + b_{t+1} \right) \prod_0^t q_s \right] = 0.$$

Households thus solve the following problem

$$\begin{aligned} & \max_{c, n, b, k} \sum_{t=0}^{\infty} \beta^t u(c_t, n_t) \\ \text{subject to} & \quad c_t - w_t(1 - \tau_t^n)n_t - T_t + k_{t+1} + q_t b_{t+1} = R_t(1 - \tau_t^k)k_t + b_t \\ & \quad \lim_{t \rightarrow \infty} \left[\left(\frac{k_{t+1}}{q_t} + b_{t+1} \right) \prod_0^t q_s \right] = 0 \\ & \quad c_t, n_t, k_{t+1} \geq 0, \text{ given } k_0 \text{ and } b_0. \end{aligned}$$

Since we are maximizing a concave function on a convex set, the following first order conditions, together with the government's budget constraint and the transversality condition, are necessary and sufficient for household optimality

$$\begin{aligned} u_n(c_t, n_t) + w_t(1 - \tau_t^n)u_c(c_t, n_t) &= 0 \\ k_{t+1}[u_c(c_t, n_t) - \beta R_{t+1}(1 - \tau_{t+1}^k)u_c(c_{t+1}, n_{t+1})] &= 0 \\ q_t u_c(c_t, n_t) - \beta u_c(c_{t+1}, n_{t+1}) &= 0. \end{aligned}$$

Furthermore, the following conditions are necessary and sufficient for firm optimality and market clearing

$$\begin{aligned} w_t &= F_n(k_t, n_t) \\ R_t &= F_k(k_t, n_t) \\ c_t + g_t + k_{t+1} &= F(k_t, n_t). \end{aligned}$$

Finally, the government has to meet the following constraints

$$g_t + b_t + T_t = w_t \tau_t^n n_t + R_t \tau_t^k k_t + q_t b_{t+1} \text{ and } T_t \geq 0.$$

Thus an allocation is attainable under commitment if and only if there are prices and policies such that all the above conditions are met.

We can drop the government's budget constraint since the resource constraint and the households' budget constraint jointly imply that the government's budget constraint is always met.

Multiplying the households' budget constraint by $u_c(c_t, n_t)$ and plugging in the first order condition leads to the following implementability condition

$$\begin{aligned} c_t u_c(c_t, n_t) + n_t u_n(c_t, n_t) + \beta u_c(c_{t+1}, n_{t+1}) \left[\frac{u_c(c_t, n_t)}{\beta u_c(c_{t+1}, n_{t+1})} k_{t+1} + b_{t+1} \right] \\ = u_c(c_t, n_t) \left[\frac{u_c(c_{t-1}, n_{t-1})}{\beta u_c(c_t, n_t)} k_t + b_t \right] + T_t u_c(c_t, n_t), \end{aligned}$$

which can replace the households' budget constraint.

The only conditions that constrain the allocations achievable under commitment are the implementability condition, the resource constraint, and the transversality condition. All the remaining conditions can be met by choosing prices and policies. Furthermore, all the initial conditions can be recovered using this characterization of the economy. Thus, an allocation is attainable under commitment if and only if it meets the following conditions

$$\begin{aligned} m(c_t, n_t) + \beta a_{t+1} &\geq a_t \\ c_t + g_t + k_{t+1} &= F(k_t, n_t) \\ \lim_{t \rightarrow \infty} [\beta^t a_{t+1}] &= 0, \end{aligned}$$

with $a_t \equiv u_c(c_t, n_t) \left[\frac{u_c(c_{t-1}, n_{t-1})}{\beta u_c(c_t, n_t)} k_t + b_t \right]$ for $t > 0$ and $a_0 \equiv u_c(c_0, n_0) [F_k(k_0, n_0) (1 - \tau_0^k) k_0 + b_0]$.

The prices and policies that ensure that the remaining conditions are met are the

following

$$\begin{aligned}
T_t &= m(c_t, n_t) + \beta a_{t+1} - a_t \\
w_t &= F_n(k_t, n_t) \\
R_t &= F_k(k_t, n_t) \\
\tau_t^n &= 1 + \frac{1}{w_t} \frac{u_n(c_t, n_t)}{u_c(c_t, n_t)} \\
\tau_{t+1}^k &= 1 - \frac{1}{\beta R_t} \frac{u_c(c_t, n_t)}{u_c(c_{t+1}, n_{t+1})}.
\end{aligned}$$

1.7.2 Proof of Lemma 2

The value of the worst sustainable equilibrium depends only on the payoff relevant state variables, which are the current stock of capital and bonds, as well the percentage of households who defaulted in the previous period.

Suppose that this is not the case and that there are two histories A and B with the same k , b , and d^c , but with $V^w(A) > V^w(B)$. Then the value of A could be decreased if all agents followed the strategies from equilibrium B , which means it could not have been the worst equilibrium to begin with. Thus, it must be true that $V(\sigma^w, f^w, \phi^w) = V^w(k, b, d^c)$.

Since $V^w(k, b, 0)$ was the worst sustainable equilibrium when no consumers defaulted in the previous period, if $d^c > 0$ and the government decides not to punish consumers, the welfare cannot be lower than $V^w(k, b, 0)$. This implies the value of the worst sustainable equilibrium does not depend on whether consumers defaulted in the previous period.

The value of the worst sustainable equilibrium is also independent of b . Since $V^w(k, 0, d^c)$ is the worst equilibrium when the government has no debt, then if the government defaults on its debt no additional punishment can be given to him, which means that $V^w(k, b, d^c) \geq V^w(k, 0, d^c)$. If $b < 0$, then $V^w(k, b, d^c) \leq V^w(k, 0, d^c)$ since the politician can achieve $V^w(k, 0, d^c)$ by making transfers to consumers and equilibrium with $b < 0$ was already the worst possible, which means that no additional loss of welfare will be possible. Furthermore, it cannot be the case that $V^w(k, b, d^c) >$

$V^w(k, 0, d^e)$ when $b > 0$ since then there would be a worst equilibrium where all households default.

Thus, the value of the worst sustainable equilibrium can only depend on the initial level of capital.

1.7.3 Proof of Lemma 3

In the worst sustainable equilibrium, households' beliefs about the following subgame are manipulated to yield the worst possible payoffs for the government. As we have seen, in each period households make their decisions according to

$$\begin{aligned} u_c(c, n)F_n(k, n)(1 - \tau^n) + u_n(c, n) &= 0 \\ k'\{u_c(c, n) - \beta u_c(c', n')F_k(k', n')(1 - \tau^{k'})\} &= 0 \\ c_t + k_{t+1} &= R_t(1 - \tau_t^k)k_t + w_t(1 - \tau_t^n)n_t + T_t \end{aligned}$$

Following Phelan and Stachetti (2001), define z' as the marginal value of capital for households tomorrow $z' \equiv u_c(c', n')F_k(k', n')(1 - \tau^{k'})$. The only way to affect households actions is by changing this value.

Let us now consider the government's problem when faced with a worst sustainable equilibrium. Let $V^w(k)$ be the expected payoff of the worst sustainable equilibrium when initial capital is k . Since the worst equilibrium the government can be given tomorrow is $V^w(k')$, then the following equilibrium is always available, which means that $V^w(k) \geq \underline{V}(k, z')$

$$\begin{aligned} \underline{V}(k, z') &= \max_{c, n, g, k'} \{u(c, n) + v(g) + \beta V^w(k')\} \\ \text{subject to} \quad m(c, n) + k'u_c(c, n) &\geq 0 \\ c + g + k' &= F(k, n) \\ k'[u_c(c, n) - \beta z'] &= 0. \end{aligned}$$

Let $z' \in [z, \bar{z}]$ be the values of z that can be sustained tomorrow. Then the worst sustainable equilibrium must be given by

$$V^w(k) = \min_{z' \in [\underline{z}, \bar{z}]} \underline{V}(k, z').$$

From government's optimality we cannot have $V^w(k) < \min \underline{V}(k, z')$.

Suppose $V^w(k) > \min \underline{V}(k, z')$. Then this could not be the worst sustainable equilibrium since a lower payoff could be reached by giving consumers expectations $z' = \arg \min \underline{V}(k, z')$.

Since $\underline{V}(k, z')$ is increasing in z' , it achieves its minima at the lower bound $\underline{z} = 0$. This means that consumers will expect capital to be taxed at confiscatory rates tomorrow, which implies that there will be no investment in capital. Furthermore, it will always be incentive compatible for the government to expropriate capital. Thus, the default equilibrium is the worst sustainable equilibrium.

1.7.4 Proof of Lemma 4

Suppose that the allocation x is the outcome of a sustainable equilibrium (σ, f, ϕ) . Consumer optimality requires that x maximizes the households utility at time zero given the policies and prices along the equilibrium path. Government optimality implies that the government must satisfy its budget constraints from time zero on. Furthermore, in a sustainable equilibrium factor prices must equal marginal productivity of factors and the resource constraint must hold along the equilibrium path. Thus, the outcome of a sustainable equilibrium x must be attainable under commitment, which means that condition (i) must hold. At any time t , after history h_{t-1} , if the government deviates from the equilibrium path, it will get a payoff higher or equal to $V^d(k_t)$. Government optimality requires that not deviating must yield a higher payoff than deviating, which means that the present value of future profits must be at least as high as $V^d(k_t)$, which means that condition (ii) must hold at every time t . Thus, if an allocation x is the outcome of a sustainable equilibrium (σ, f, ϕ) it must meet conditions (i) and (ii).

Suppose now that an allocation x meets conditions (i) and (ii). Let π and ϕ be the policies and prices that implement this allocation under commitment. Consider

the following strategy for households: as long as the government's action is according to π , choose allocation x ; if the government deviates, follow the default equilibrium strategy. Likewise, consider the government's strategy where it acts according to π along the equilibrium path and plays the default strategies off equilibrium. Finally, consider factor prices that are equal to the marginal productivity of each factor of production for every possible history h_t . We will show that this is a sustainable equilibrium. First, consider histories where there have been no deviations until time t . Since x is attainable under equilibrium, and along the equilibrium path households will expect to face policies π and prices ϕ , this means that the continuation of x must be optimal for consumers. The government, on the other hand, can choose to deviate, in which case it would get $V^d(k_t)$, or it can follow the equilibrium path. Since condition (ii) ensures that the payoff along the equilibrium path is always higher than $V^d(k_t)$, then it is always incentive compatible for the government not to deviate. Now, consider histories where there has been a deviation before time t . Our strategy has specified that in this case both households and the government will play a default equilibrium, which we have shown to be sustainable. Thus, the specified set of strategies is a sustainable equilibrium that leads to outcome x .

1.7.5 Equivalence of Sequence and Recursive Approaches

The Lagrangean for the sequence problem to find the best sustainable equilibrium can be written in the following way, where $\hat{\mu}_t$ is the multiplier on the implementability condition, $\hat{\rho}_t$ is the multiplier on the resource constraint, and $\hat{\gamma}_t$ is the multiplier on incentive compatibility condition

$$n = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t) + \hat{\mu}_t [m(c_t, n_t) + \beta a_{t+1} - a_t] \\ - \hat{\rho}_t [c_t + g_t + k_{t+1} - F(k_t, n_t)] \end{array} \right\} + \sum_{i=1}^{\infty} \beta^i \hat{\gamma}_i \left[\sum_{t=i}^{\infty} \beta^{t-i} u(c_t, n_t) - V^d(k_i) \right].$$

The first order conditions for this problem are

$$\begin{aligned}
u_{c_t}(1 + \sum_{i=1}^t \widehat{\gamma}_i) + \widehat{\mu}_t m_{c_t} &= \widehat{\rho}_t \\
u_{n_t}(1 + \sum_{i=1}^t \widehat{\gamma}_i) + \widehat{\mu}_t m_{n_t} &= -\widehat{\rho}_t F_{n_t} \\
\beta \widehat{\rho}_{t+1} F_{k_{t+1}} - \beta \widehat{\gamma}_{t+1} V_k^d(k_{t+1}) &= \widehat{\rho}_t \\
\widehat{\mu}_{t+1} &= \widehat{\mu}_t.
\end{aligned}$$

The first order conditions for the recursive problem in section 1.4.2 are

$$\begin{aligned}
u_c + \mu m_c &= \rho \\
u_n + \mu m_n &= -\rho F_n \\
v_g &= \rho \\
V_a(a', k')(1 + \gamma) &= -\mu \\
V_k(a', k') &= \rho/\beta + \gamma[V_k^d(k') - V_k(a', k')]
\end{aligned}$$

and the envelope conditions for a and k are

$$\begin{aligned}
V_a(a, k) &= -\mu \\
V_k(a, k) &= \rho F_k(k, n).
\end{aligned}$$

It is easy to see that the equilibrium conditions for the two problems lead to the same allocations for c , n , k , and a for as long as the multipliers for the constraints in the recursive formulation μ , ρ , and γ have the following relationship with the

multipliers in the sequence approach

$$\begin{aligned}\mu &= \frac{\hat{\mu}_t}{1 + \sum^t \hat{\gamma}_i} \\ \rho &= \frac{\hat{\rho}_t}{1 + \sum^t \hat{\gamma}_i} \\ 1 + \gamma &= \frac{1 + \sum^{t+1} \hat{\gamma}_i}{1 + \sum^t \hat{\gamma}_i}.\end{aligned}$$

Furthermore, the following transversality condition must be met in both formulations

$$\lim_{t \rightarrow \infty} \beta^t V(k_t, a_t) = 0.$$

Since the best sustainable equilibrium converges to a steady state, which has positive and finite allocations, the transversality condition is met.

1.7.6 Impossibility of non zero Capital Taxes in Steady State

We will now show that any steady state of the best sustainable equilibrium must have zero capital taxes without imposing concavity of $V(k_t, a_t)$.

The idea is that if commitment were binding in the long run, then the implementability condition would stop binding, and labor taxes would converge to zero. But this can not be an optimal steady state, since it is optimal to take a deviation where labor taxes are increased marginally (with zero first order cost) to increase government assets and make the incentive compatibility constraint less binding (which has a positive first order effect).

More formally, assume that there is a steady state where capital taxes are different from zero. Let c^{ss} , n^{ss} , g^{ss} , a^{ss} and k^{ss} be the allocations in this steady state. In order to meet first order conditions, these steady states must also have constant multipliers ρ^{ss} , μ^{ss} , γ^{ss} . In steady state, the optimality condition for capital becomes

$$\rho^{ss} F_k(k^{ss}, n^{ss}) = \rho^{ss} / \beta + \gamma^{ss} [V_k^d(k^{ss}) - \rho^{ss} F_k(k^{ss}, n^{ss})].$$

Given this we can only have capital taxes different from one and $\rho F_k(k^{ss}, n^{ss}) \neq 1$

if $\gamma^{ss} > 0$.

The optimality condition for a implies that

$$\mu^{ss}(1 + \gamma^{ss}) = \mu^{ss}.$$

Thus, when $\gamma^{ss} > 0$ we must have $\mu^{ss} = 0$.

But then the first order conditions for consumption, labor and public spending take the following form

$$\begin{aligned} u_c(c^{ss}, n^{ss}) &= \rho^{ss} \\ u_n(c^{ss}, n^{ss}) &= -\rho^{ss} F_n(k^{ss}, n^{ss}) \\ v'(g^{ss}) &= \rho^{ss}. \end{aligned}$$

Furthermore, it must be true that

$$V(k^{ss}, a^{ss}) = V^d(k^{ss}) < V(k^{ss}, \underline{a}(k^{ss})) = V^{fb}(k^{ss}).$$

Now let us consider a departure from this steady state that will lead to higher welfare than our original candidate for a steady state, thus implying that it could not be an optimal solution to begin with. The departure is as follows. In an initial period (let us call it period 0), starting from our initial steady state, we will choose $a_1 = a^{ss} - \Delta a$ instead of a^{ss} and $k_1 = k^{ss}$. The new levels of consumption, labor and public spending are the solution to the following problem

$$\begin{aligned} [c_0, n_0, g_0] &= \arg \max_{c, n, g} [u(c, n) + v(g)] \\ \text{subject to} & \quad w(c, n) + \beta(a^{ss} - \Delta a) \geq a^{ss} \\ & \quad c + g + k^{ss} = F(k^{ss}, n). \end{aligned}$$

First, let us check that this deviation is feasible. Clearly the implementability condition and the resource constraint must be met, by construction of the previous

problem. The government's incentive compatibility condition now becomes

$$V(k^{ss}, a^{ss} - \Delta a) \geq V^d(k^{ss}).$$

Since $V(k, a)$ cannot be increasing in a this condition must be met since $V(k^{ss}, a^{ss}) = V^d(k^{ss})$ in the original steady state.

Now we will show that taking this deviation has a zero cost up to a first order approximation. The new flow utility in period zero is given by

$$\begin{aligned} W(\Delta a | k^{ss}, a^{ss}) &= \max_{c, n, g} [u(c, n) + v(g)] \\ \text{subject to} \quad &w(c, n) + \beta(a^{ss} - \Delta a) \geq a^{ss} \\ &c + g + k^{ss} = F(k^{ss}, n). \end{aligned}$$

We can write a first order Taylor approximation of this expression in the following way

$$W(\Delta a) = W(0) + \Delta a W'(0).$$

Given that the implementability is not binding when $\Delta a = 0$, then $W(\Delta a) = W(0)$, which means that the cost of this deviation is zero to a first order approximation.

Let us now see if there is any benefit to it. The deviation proceeds as follows. From period 1 to period $T - 1$ (which will be defined shortly), the chosen allocations will be

$$\begin{aligned} c_t &= c^{ss}, n_t = n^{ss}, g_t = g^{ss}, k_t = k^{ss} \text{ for } t = 1, \dots, T - 1 \\ a_t &= a^{ss} - \beta^{1-t} \Delta a \text{ for } t = 1, \dots, T. \end{aligned}$$

Let us start by checking that the new plan is feasible from periods 1 to T . First, the resource constraint must be met since the allocations in this constraint are the same as in the initial steady state. The incentive compatibility must also hold since $V(k, a)$ cannot be increasing in a . Finally, the implementability condition will be met since

$w(c^{ss}, n^{ss}) \geq a^{ss}(1 - \beta)$, which implies that $w(c, n) + \beta(a^{ss} - \beta^{-t} \Delta a) \geq a^{ss} - \beta^{1-t} \Delta a$.

Let T be defined by the following condition

$$\Delta a \frac{1}{\beta^T} = a^{ss} - \underline{a}(k^{ss}).$$

This means that in period T we will reach $a_t = \underline{a}(k^{ss})$ and will be able to increase $V(k^{ss}, \underline{a}(k^{ss})) > V(k^{ss}, a^{ss})$ by switching to the non constrained solution. Furthermore, this will lead to a first order positive welfare increase in the initial period

$$B(\Delta a) = \beta^T [V(k^{ss}, \underline{a}(k^{ss})) - V(k^{ss}, a^{ss})] = \frac{V(k^{ss}, \underline{a}(k^{ss})) - V(k^{ss}, a^{ss})}{a^{ss} - \underline{a}(k^{ss})} \Delta a.$$

Thus, the net benefit of our deviation is strictly positive for a small enough change in a , which means that the initial candidate for a steady state was not optimal.

Chapter 2

Education and Optimal Taxation

The optimal income taxation problem studied by Mirrlees takes individual ability as given. However, ability is to a large extent determined by the amount of education acquired by workers, which is potentially influenced by income taxes. Thus, the optimal income tax schedule should take this effect into consideration. This chapter models the optimal taxation problem when education is endogenous and observable. In this context, it is optimal to subsidize education in order to offset the distortion caused by the income tax. If the cost of education is purely monetary, the education subsidy exactly offsets the income tax, so that the choice of education is not distorted by income taxes. On the other hand, if the cost of education is an opportunity cost of time, then the subsidy on education only partially offsets the income tax and the choice of education is still distorted by taxation.

2.1 Introduction

Mirrlees (1971) was the first to analyze the income taxation problem allowing a general tax function and considering a continuum of agents that differ in their innate ability to produce output. Since this ability is private information, the government must create a distortion in labor supply in order to do some redistribution. The optimal tax policy results from the optimal trade-off between efficiency and income redistribution. Later work by Mirrlees (1986) extends his analysis to include a vector of decision variables

and multiple individual attributes. In the model analyzed in this chapter there are two decision variables for individuals (choice of hours worked and level of education), but only one-dimensional heterogeneity, which means it is a particular case of the setup analyzed by Mirrlees. The contribution of this model is thus to highlight the particular implications that Mirrlees' results have in the presence of observable education.

The interaction between education and redistributive policy was first analyzed by Hare and Ulph (1979), but they ignore the choice of labor when education is observable, so in their model first best is achievable. Broadway, Marceau and Marchand (1996) focus on the time inconsistency of the optimal policy when education is endogenous but unobservable. Diamond and Mirrlees (2003) show that the presence of endogenous education need not make it optimal to reduce income taxation if the income derivative of education is high enough. Brett and Weymark (2003) look at what the optimal income tax looks like in a context where education is fully subsidized. Gottlieb and Moreira (2003) focus on the regressivity of educational subsidies (since children from wealthier families tend to receive more education than those from poor families). Costa and Maestri (2004) analyze interactions between income tax and education in a context where all individuals are identical at the time of their educational choice and their type is only revealed afterwards. Kapicka (2004) looks at what optimal income tax looks like in a dynamic model with (unobservable) human capital; using a model with individual heterogeneity (building on the Mirrlees model), he finds that in the long run human capital makes less redistribution optimal; he also runs some numerical simulations for the US economy that suggest that, in the presence of human capital, optimal tax rates will be 10-16% lower.

This chapter derives the optimal tax schedule in the presence of observable education choices (and unobservable labor). It is common in the income taxation literature to assume labor (or its cost for individuals) is not observable, even though hours of work could conceivably be observed if the planner so desired to do. One first reason is that hours of work do not necessarily reflect effort, as the vast contracting literature on providing adequate incentives at the firm level shows. If not even the employer can observe the actual input, then it must be even harder for the government. Fur-

thermore, even if effort is observable at the firm level, if taxes are diminishing in the number of hours worked, this provides an incentive for workers to stay additional hours on the job doing private tasks that might otherwise be done more efficiently at home.

Education, on the other hand, might be more easily measured. One major reason is that a large part of its cost is indeed monetary and easily measurable. Furthermore, this cost is a priori the same for all individuals. Naturally one might imagine that some people have a smaller cost of learning, but even if it is relevant it might be offset by being strongly correlated with higher foregone wages, which in turn increase the cost of education.

It is also assumed that agents have full information about their type and make their education and labor choices for a given tax schedule, which is contingent on the level of education and labor income. All agents are assumed to have zero initial wealth, so all inequality comes from different ability levels. Furthermore, there is an implicit assumption that agents can freely borrow against future wealth and have the same discount rate as the government.

Under these conditions, if the disutility of labor is not significantly affected by other variables in the utility function (namely, if the level of income and education does not affect the disutility of labor) then it is optimal not to distort the choice of education. This can be achieved by giving an educational subsidy that offsets the distortion that income tax has on the choice of education. On the other hand, if for a higher level of education (or income) the cost of labor increases, then the direct subsidy on education should be lower (or even negative in very extreme cases). Finally, it might be optimal give a subsidy to education that is even larger than the tax on income if either education or wealth make labor time more agreeable.

Independent work by Bovenberg and Jacobs (2005) finds similar conclusions to this chapter.

2.2 The model

Agents have a utility function $U(y, e, l)$ over net income y , the level of education e and labor supply l . Utility is non-decreasing and concave in $x \equiv (y, -e^2, -l^2)$. It needs to be more concave than it is usual to assume since the production function will exhibit increasing returns to scale, which need to be offset to guarantee the problem can be solved using a first order approach. Assume also that the following *inada* conditions hold: $U_y(0, e, l) = \infty$, $U_y(\infty, e, l) = 0$, $U_e(y, 0, l) = 0$, $U_e(y, \infty, l) = -\infty$, $U_l(y, e, 0) = 0$ and $U_l(y, e, \infty) = -\infty$. This function is defined for non negative values of y , e and l .

Consider a static decision problem where agents choose education and labor in the same period as they consume. This formulation can be seen as a reduced form of a dynamic problem where the agent maximizes lifetime utility.

Individuals can have different types that are characterized by their ability θ , which is distributed between $\underline{\theta}$ and $\bar{\theta}$ with *cdf* $F(\theta)$ and *pdf* $f(\theta)$.

Each individual can choose a level of education (measured in years of schooling) which enhances productivity proportionately to individuals' initial ability. Furthermore, each agent's education only influences his own productivity (so there is no externality of education on the overall productivity of the economy), so that labor income for a given agent is given by

$$w = \theta el.$$

This complementarity between education and ability is crucial for the results of this paper, but it seems to be a reasonable one. With complementarity between ability and education it will be more productive for relatively higher types to invest in education. Regev (2005) argues that one possible reason why lower ability workers might benefit more from acquiring education is that they have a lower cost of foregone earnings. Even though this possibility is allowed by this model (by allowing education to crowd out labor in the utility function), it does not change the results. Furthermore, productivity does not have to depend linearly on education. All the results follow

through if we replace e by a strictly increasing function of education $\phi(e)$ such that labor income would now be given $w = \theta\phi(e)l$ (as long as the concavity assumptions for the utility function as changed accordingly to ensure that the problem remains convex).

The government maximizes a social welfare function given by

$$W = \int_{\underline{\theta}}^{\bar{\theta}} G(u(\theta))dF(\theta)$$

and has to finance unproductive expenditures in the amount E . For this, it can use taxes which are potentially a non-linear function of observable labor income and education: $T(w, e)$. The government budget constraint is thus given by

$$\int_{\underline{\theta}}^{\bar{\theta}} T(w_{\theta}, e_{\theta})dF(\theta) \geq E.$$

The net income available for agent's consumption is given by total income minus taxes

$$y = w - T(w, e).$$

In addition to its budget constraint, the government also has to consider the incentive compatibility conditions for the agents in order to guarantee that they do not want to imitate another type

$$\begin{aligned} & U(\theta e_{\theta} l_{\theta} - T(\theta e_{\theta} l_{\theta}, e_{\theta}), e_{\theta}, l_{\theta}) \\ & \geq U(\theta' e_{\theta'} l_{\theta'} - T(\theta' e_{\theta'} l_{\theta'}, e_{\theta'}), e_{\theta'}, \frac{\theta' e_{\theta'} l_{\theta'}}{\theta}) \\ & \text{for all } \theta \text{ and } \theta'. \end{aligned}$$

Since this condition is not very manageable, replace it for the first order condition of the consumer's problem.

Given the assumptions of the model, the first order conditions characterize the optimal decision for the agents when taxes are zero. However, this may not be the case for the optimal tax schedule since its functional form is endogenous, which may

lead to a non-convex budget set when marginal income tax rates are declining. This chapter will abstract from this problem and continue the analysis using first order conditions, which can be written as

$$\begin{aligned} U_y[\theta l [1 - T_w(\theta e l, e)] - T_e(\theta e l, e)] + U_e &= 0 \\ U_y \theta e [1 - T_w(\theta e l, e)] + U_l &= 0. \end{aligned}$$

Using these conditions as a constraint for the planner's problem, we can derive the first order conditions for the social optimum. These conditions are a particular case of the optimality conditions for the problem studied by Mirrlees (1986) for non-linear taxation in a one-dimensional population¹. The simplified conditions are

$$\frac{T_w}{1 - T_w} = \frac{\sigma_\theta}{\theta \mu_\theta} A, \text{ where } A = 1 + l_\theta \frac{U_{ll}}{U_l} + \theta e_\theta l_\theta \frac{U_{ly}}{U_l}$$

$$T_e = T_w \frac{U_e}{U_y} \frac{A - B}{A}, \text{ where } B = l_\theta \frac{U_{le}}{U_e} + \theta l_\theta^2 \frac{U_{ly}}{U_e}.$$

If U_{le} and U_{ly} are small then $T_e \simeq T_w \frac{U_e}{U_y}$. If we replace this expression in the consumer's first order condition for the choice of education, we get $U_y \theta l_\theta + U_e \simeq 0$ which means that the education choice will not be distorted by taxes. Note that this does not mean that the level of education will be the same as if there were no taxes since taxes reduce labor supply, which in turn reduces the benefit of education, making it optimal to invest less in it.

However, it is possible that these cross derivatives are not negligible. In the next section we see what happens if education crowds out work hours.

2.2.1 Time Cost of Education

Assume that utility is quasi-linear. This particular case was studied by Diamond (1998) so it will give us a useful benchmark for our results.

¹For a direct derivation see the appendix. σ_θ is the multiplier on the agent's incentive compatibility constraint and μ_θ is the multiplier on the constraint that determines the level of utility a given type attains given the amount of taxes it pays.

Utility is now linear in income, but there is a convex cost of spending time either acquiring education or working. We can think of the cost of education as being the opportunity cost of the time that could have been spent working

$$U(y, e, l) = y - v(e + l).$$

Since the labor and education are substitutes in the utility function, this means that whenever education is higher, this will not only increase the marginal disutility of education, but also the marginal disutility of labor.

In this case, the first order conditions for the consumer's problem become

$$\begin{aligned}\theta l [1 - T_w(\theta e l, e)] &= v'(e + l) + T_e(\theta e l, e) \\ \theta e [1 - T_w(\theta e l, e)] &= v'(e + l).\end{aligned}$$

The optimality conditions for the allocations are

$$\begin{aligned}\sigma'_\theta &= f(\theta)[\rho - G'(u(\theta))] \\ f(\theta)\rho[\theta l_\theta - v'(e_\theta + l_\theta)] &= \frac{\sigma_\theta l_\theta}{\theta} v''(e_\theta + l_\theta) \\ f(\theta)\rho[\theta e_\theta - v'(e_\theta + l_\theta)] &= \frac{\sigma_\theta l_\theta}{\theta} \left[\frac{v'(e_\theta + l_\theta)}{l_\theta} + v''(e_\theta + l_\theta) \right].\end{aligned}$$

Using the fact that $\sigma(\bar{\theta}) = 0$, we can find $\sigma(\theta)$ by integrating $\sigma'(\theta)$ and using the consumer's first order conditions to recover the tax rates, we can characterize the optimal tax schedule

$$\frac{T_w}{1 - T_w} = \frac{\int_{\bar{\theta}}^{\theta} (\rho - G') dF}{\rho(1 - F)} \frac{1 - F}{\theta f} \eta_l,$$

$$T_e = -\frac{v' T_w}{\eta_l}.$$

The first expression characterizes the optimal income tax schedule² and replicates the result of Diamond (1998) for a given level of education, which suggests that the

² $\eta_l \equiv 1 + \varepsilon_l^{-1}$, where ε_l is the elasticity of labor supply which is given by $\varepsilon_l = v'/(v''l)$.

presence of endogenous education does not influence optimal income taxes (except by changing the underlying education level).

Up until now we have been talking about *gross* taxes on income and education, which are given by T_w and T_e , respectively. Let us now consider the *net* tax on education, which I will define as the direct gross tax on education plus the tax levied on education through income taxes and which is given by $T_e + \theta l T_w$. Plugging into the consumer's first order condition, we can see that the net tax on education actually gives us the direct distortion that taxes have on the choice of education, which is given by $\theta l - v'$.

Proposition 4 *With time cost of education the optimal net tax on education is positive.*

To see this replace the optimal education tax into the consumer's first order condition for education to get

$$\frac{\theta y}{v'} = \frac{1 - T_w/\eta_l}{1 - T_w} > 1.$$

Since the left hand side is increasing in the level of education, this will distort education downwards. This happens because by reducing the amount of education we can reduce the distortion in labor supply, since more education makes labor more expensive.

Note that despite the fact that *net* taxes on education are positive, the optimal *gross* tax on education is actually negative (which means there is an education subsidy); it is just not large enough to offset the taxation levied on education through income taxes.

2.2.2 Monetary Cost of Education

Assume now that utility is still quasi-linear, but it is additively separable in education and labor in the following way

$$U(y, e, l) = y - c(e) - v(l).$$

In this formulation, we can interpret $c(e)$ as a monetary cost of education that does not affect utility in any way other than reducing the amount of income available for consumption.

The first order conditions for the consumer's problem are now given by

$$\begin{aligned}\theta l [1 - T_w(\theta ey, e)] &= T_e(\theta ey, e) + c'(e) \\ \theta e [1 - T_w(\theta ey, e)] &= v'(l).\end{aligned}$$

The optimality conditions for the allocations are

$$\begin{aligned}\sigma'_\theta &= f(\theta)[\rho - G'(u(\theta))] \\ \theta l_\theta &= c'(e_\theta) \\ f(\theta)\rho[\theta e_\theta - v'(l_\theta)] &= \frac{\sigma_\theta l_\theta}{\theta} \left[\frac{v'(l_\theta)}{l_\theta} + v''(l_\theta) \right]\end{aligned}$$

and the optimal tax schedule is given by

$$\begin{aligned}\frac{T_w}{1 - T_w} &= \frac{\int_{\bar{\theta}}^{\bar{\theta}} (\rho - G') dF}{\rho(1 - F)} \frac{1 - F}{\theta f} \eta_l \\ T_e &= -c' T_w.\end{aligned}$$

As in the case with time cost of education, the first expression characterizes the optimal income tax schedule and is identical to the expression reached by Diamond (1998).

Proposition 5 *With monetary cost of education the optimal net tax on education is zero.*

To see this, replace the optimal education tax into the consumer's first order condition for education to get

$$\theta y = c',$$

which means that it is optimal not to distort the choice of education since for each type, we are subsidizing the cost of education at the same marginal rate at which

we are marginally taxing labor income. At first this result may seem odd since it seems that a perturbation of education around its first best level would have only a negligible effect on utility while allowing us to keep more subsidy money, but this is not the case because the reduction in education would also lead to a decrease in collected taxes (since we would be in a region where the marginal cost of education is lower than the marginal benefit, then paying T' of that cost as an education subsidy and receiving T' of the benefit as taxes would actually increase government revenue).

Proposition 6 *In the model with monetary cost of education the optimal solution is the same for a generic function $T(w, e)$ or for a function that depends only on income net of educational expenses $t(w - c(e))$ as long as $w - c(e)$ is strictly increasing in θ .*

It is trivial to go from the particular functional form to the more general case, so the proof focuses on the opposite direction.

Choose $t'(w(\theta) - c(e(\theta))) = T_w(w(\theta), e(\theta))$. We can do this since $w - c(e)$ is strictly increasing in θ . Note that $\frac{dt}{de} = -c'e' = -c'T_w = T_e$ for this particular formulation of the utility function. This is a crucial condition for this proof to work because unless this is true we will not be able to match the consumer's first order conditions.

$$\text{Set } t(w(\underline{\theta}) - c(e(\underline{\theta}))) = T(w(\underline{\theta}), e(\underline{\theta})).$$

Furthermore, define $\tau(\theta) \equiv T(w(\theta), e(\theta))$ and $\hat{\tau}(\theta) \equiv t(w(\theta) - c(e(\theta)))$. Note that $\tau'(\theta) = T_w w' + T_e e' = T_w w' - c' T_w e' = t'(w' - c'e') = \hat{\tau}'(\theta)$. Since we defined $\tau(\underline{\theta}) = \tau'(\underline{\theta})$, this implies that $\tau(\theta) = \hat{\tau}(\theta)$ for every θ , which is the same as saying that $t(w(\theta) - c(e(\theta))) = T(w(\theta), e(\theta))$ (note that this does not mean that $t(w - c(e)) = T(w, e)$ in general).

But now it is clear to see that the allocation chosen by consumers will be the same in both cases since both the first order conditions and the budget constraint are the same. Furthermore, the government will be able to collect exactly the same amount of taxes, which means that we can reach the best allocation using a generic tax function that depends on labor income and education even if we restrict ourselves to a tax function that depends only on labor income net of educational expenses.

2.2.3 Time and Monetary Cost of Education

Assume now that utility remains quasi-linear, but has *both* a time and a monetary cost associated to education

$$U(y, e, l) = y - c(e) - v(e + l).$$

This case combines the previous two and the optimal tax schedule can now be written as

$$\frac{T_w}{1 - T_w} = \frac{\int_{\theta}^{\bar{\theta}} (\rho - G') dF}{\rho(1 - F)} \frac{1 - F}{\theta f} \eta_i$$

$$T_e = -c' T_w - \frac{v' T_w}{\eta_i}.$$

In this case, it is optimal to subsidize the monetary cost of education at the same rate as we are taxing income and subsidize the time cost at a lower rate. Thus, as long as the time cost component is relevant the *net* tax on education will remain positive.

2.3 Comparing Different Models of Education

So far it has been assumed that education choice is observable by the social planner. This section analyzes what happens to the results of the model if we change this assumption.

It starts by assuming that education is not observable, but is still affected by the income tax, so that the social planner can take this additional distortionary margin into account when deciding on the optimal income tax schedule. Although the case with unobservable education has already been studied in the literature, it is useful to rederive it under my previous setup in order to be able to pinpoint the fundamental differences between the two cases more easily.

Next, the equilibrium is derived for an economy where education is still affected by income taxes, but the planner does not take this effect into account when choosing the income taxes. In this setup, the planner takes the distribution of ability after

education is chosen as given and chooses the income tax optimally for that distribution without considering that the choice of education is actually endogenous.

The section finishes with a simulation of the distributions of skills for the US economy following the procedure of Diamond (1998) and then giving a rough estimate of what optimal taxes might look like under the three alternative models.

Note that there is a trivial welfare ordering between these three models. The economy with observable education achieves (weakly) higher welfare than the economy with unobservable education (since the social planner can always choose a tax schedule that is independent of education if it is optimal to do so). Furthermore, if education is unobservable, the economy where the planner is aware of the effect of income taxes on education always has a (weakly) higher welfare than the economy with the myopic planner since the allocation in the later economy can always be achieved by choosing the same income tax schedule as the myopic planner would.

Throughout this section it is assumed that the utility function is quasi-linear and additively separable in labor and education

$$U(y, e, l) = y - c(e) - v(l).$$

2.3.1 Unobservable Education

If education is not observable, the tax function can no longer be contingent on education, which means it can only depend on each individual's observed level of income: $T(w)$. The first order conditions for the consumer's choice now become

$$\begin{aligned}\theta l [1 - T_w(\theta e y)] &= c'(e) \\ \theta e [1 - T_w(\theta e y)] &= v'(l).\end{aligned}$$

These restrictions on the planner's problem can be written as the following enve-

lope conditions

$$\begin{aligned}u'(\theta) &= v'(l_\theta)l_\theta/\theta \\u'(\theta) &= c'(e_\theta)e_\theta/\theta,\end{aligned}$$

where $u(\theta) = \theta e_\theta l_\theta - T(\theta e_\theta l_\theta) - c(e_\theta) - v(l_\theta)$.

The Hamiltonian for this problem is

$$H = f(\theta)\{G(u(\theta)) + \rho[\theta e_\theta l_\theta - u(\theta) - c(e_\theta) - v(l_\theta)]\} + \sigma_\theta \frac{v'(l_\theta)l_\theta}{\theta} + \mu_\theta \frac{c'(e_\theta)e_\theta}{\theta}$$

and the optimality conditions for the allocations are

$$\begin{aligned}\sigma'_\theta + \mu'_\theta &= f(\theta)[\rho - G'(u(\theta))] \\f(\theta)\rho[\theta l_\theta - c'(e_\theta)] &= \mu_\theta e_\theta \left[\frac{c'(e_\theta)}{e_\theta} + c''(e_\theta) \right] / \theta \\f(\theta)\rho[\theta e_\theta - v'(l_\theta)] &= \sigma_\theta l_\theta \left[\frac{v'(l_\theta)}{l_\theta} + v''(l_\theta) \right] / \theta.\end{aligned}$$

Plugging in the consumer's first order conditions, we can find the optimal tax schedule³ which is given by

$$\frac{T_w}{1 - T_w} = \frac{\int_{\bar{\theta}}^{\bar{\theta}} (\rho - G') dF}{\rho(1 - F)} \frac{1 - F}{\theta f} \frac{\eta_l \eta_e}{\eta_l + \eta_e}.$$

Remember that when education was observable (or when there was no endogenous choice of education) the last term on this equation was η_l instead of $\frac{\eta_l \eta_e}{\eta_l + \eta_e}$, which means that it is smaller (at least for the same values of l) when we have an endogenous non-observable choice of education which makes income taxes more distortionary, thus making it optimal to reduce income taxes. Of course this is only a partial analysis since the overall effect also has to account for the equilibrium effect in the remaining variables.

³ η_l is the same as before and $\eta_e \equiv 1 + \varepsilon_e^{-1}$, where ε_e is the elasticity of education which is given by $\varepsilon_e = c'/(c''e)$.

2.3.2 Myopic Social Planner

Let us now consider the case when the planner only looks at the distribution of total ability after education is chosen. In this setup it does not make sense to have observable education, since this would reveal initial ability, which would make the problem uninteresting. Thus, the tax function still depends only on observed income.

When the social planner takes education as given even though it is endogenous, the consumer's first order conditions that determine the choice of labor and education are still given by

$$\begin{aligned}\theta l [1 - T_w(\theta ey)] &= c'(e) \\ \theta e [1 - T_w(\theta ey)] &= v'(l).\end{aligned}$$

However, the planner only takes the second condition into account when choosing its tax schedule. As before, this second condition can be written as

$$u'(\theta) = v'(l_\theta)l_\theta/\theta,$$

where $u(\theta) = \theta e_\theta l_\theta - T(\theta e_\theta l_\theta) - c(e_\theta) - v(l_\theta)$.

The Hamiltonian for this problem is

$$H = f(\theta)\{G(u(\theta)) + \rho[\theta e_\theta l_\theta - u(\theta) - c(e_\theta) - v(l_\theta)]\} + \sigma_\theta \frac{v'(l_\theta)l_\theta}{\theta}$$

and the optimality conditions for the allocations are

$$\begin{aligned}\sigma'_\theta &= f(\theta)[\rho - G'(u(\theta))] \\ f(\theta)\rho[\theta e_\theta - v'(l_\theta)] &= \sigma_\theta l_\theta \left[\frac{v'(l_\theta)}{l_\theta} + v''(l_\theta) \right] / \theta.\end{aligned}$$

Plugging in the consumer's first order conditions, we can find the optimal tax

schedule which is given by

$$\frac{T_w}{1 - T_w} = \frac{\int_{\theta}^{\bar{\theta}} (\rho - G') dF}{\rho(1 - F)} \frac{1 - F}{\theta f} \eta.$$

Note that this expression is exactly the same as in the case when education is observable, but there is no additional subsidy on education to offset the distortion caused by income taxes.

2.3.3 Simulation of Optimal Taxes for Different Models of Education

Following the procedure used in Diamond (1998) to get a rough idea of the skill distribution in the US, this section uses the 2003 CPS data available through IPUMS-CPS to calculate an implied average wage dividing yearly wages by weeks worked and typical hours per week. Earners of both sexes are considered, as long as their average hourly wage is above \$1 and individuals usually worked more than ten hours per week on more than ten weeks per year. The log of wage is regressed on an age polynomial, gender, marital status and education⁴ and then take the log of residuals to be my ability component. The resulting distribution is plotted below.

The relative ability of individuals is represented in the horizontal axes of Figure 2-1, with 1.0 being the average ability. We can see that the density peaks at about 75% of the mean ability. To better analyze the tail behavior of the distribution, we must find out how the ratio $[1 - F(\theta)]/f(\theta)$ evolves, which is shown in Figure 2. As is the case in Diamond (1998), we can see that $[1 - F(\theta)]/\theta f(\theta)$ stays roughly constant above the mean ability, which is consistent with a Pareto distribution.

Given this distribution of skills, we can now simulate what the optimal tax sched-

⁴Standard OLS is used to estimate the impact of education on hourly wage, which is arguably not the best procedure since the choice of education is itself affected by initial ability, which would bias the estimate upward. Many studies using instrumental variables have however not found smaller effects, which seems to suggest that both estimates are biased upward (see Card (2001) for a comparative analysis of different studies). On the other hand, we are only measuring quantity of education and not quality, which might understate its effect.

Using different coefficients for education or leaving it altogether out of the regression did not change the features of the estimated distribution of abilities.

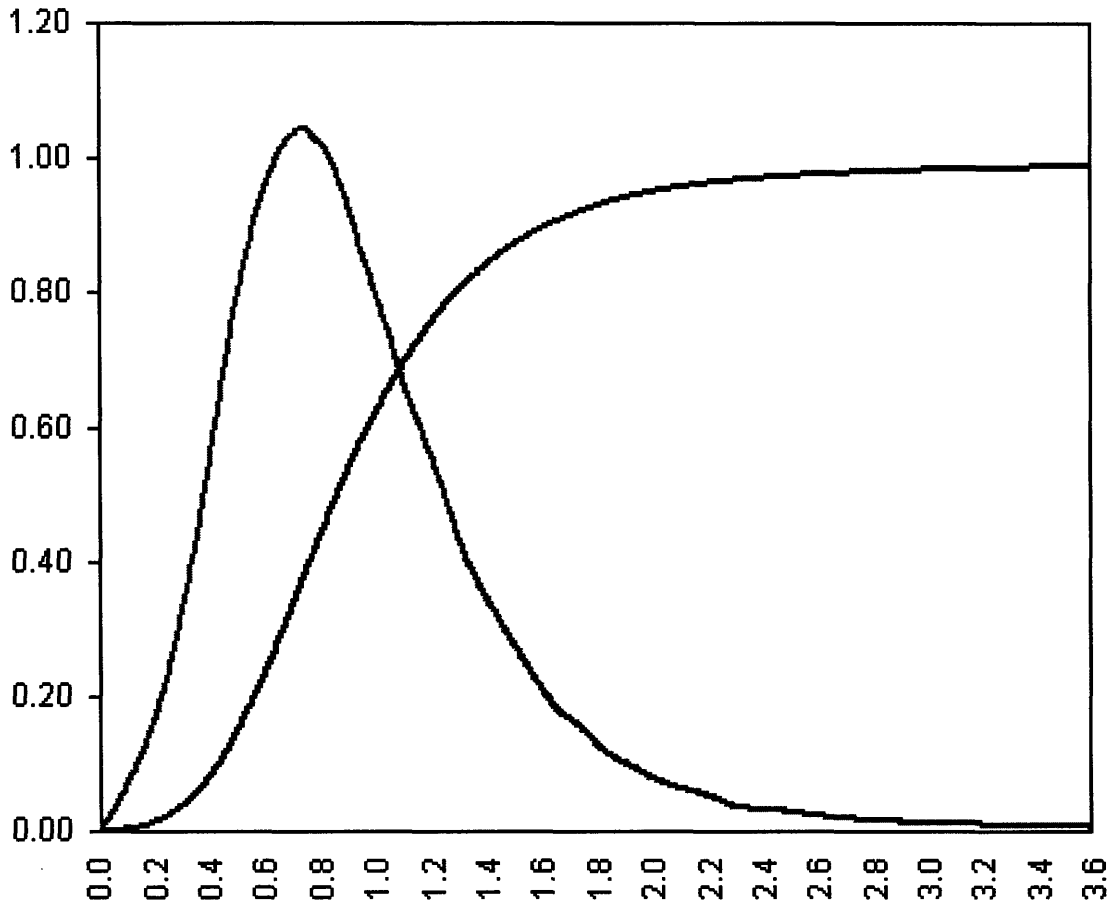


Figure 2-1: Estimated Density and Cumulative Distribution of Abilities

ule might look like. Note that this is not meant to be an exact policy prescription for what the income tax should be, since some additional non-trivial assumptions need to be made. Assume a constant elasticity of labor supply equal to $\varepsilon_l = 0.5$. Likewise, for education assume $\varepsilon_e = 0.5$. The social welfare function is $G(U) = U^{1-\sigma}/(1-\sigma)$, with $\sigma = 2$ and the government needs to collect no revenue: $E = 0$.

Following Diamond (1998), if we assume ability follows a Pareto distribution and the elasticity of labor supply is constant, we can write the optimal tax for the case when education is observable or the planner is myopic as

$$\frac{T_w}{1 - T_w} = \frac{B(\theta)}{a} \eta,$$

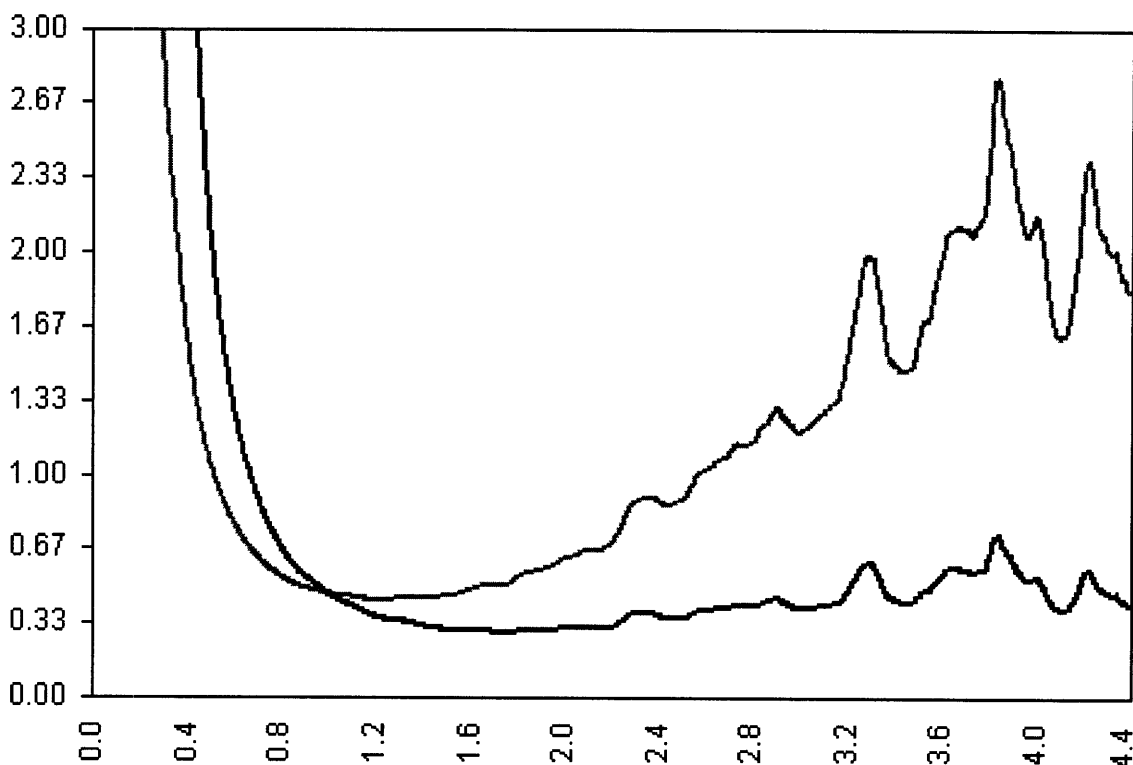


Figure 2-2: Estimated Ratios $\frac{1-F(\theta)}{f(\theta)}$ in grey and $\frac{1-F(\theta)}{\theta f(\theta)}$ in black

where $B(\theta) = \frac{\int_{\theta}^{\infty} (\rho - G') dF}{\rho(1-F)}$ and $a = \frac{\theta f(\theta)}{1-F(\theta)}$ is the coefficient of the Pareto distribution. For the chosen social welfare function, G' converges to zero as ability rises without limit, which means that $B(\theta)$ will converge to one. Furthermore, Figure 2-2 seems to suggest that that a is approximately 3 above the mean ability and for an elasticity of labor supply of 0.5 we have $\eta_l = 3$, which means that optimal T_w will converge to approximately 0.5⁵.

If education is unobservable, optimal taxes under a Pareto distribution and constant ε_l and ε_e will instead be given by

$$\frac{T_w}{1 - T_w} = \frac{B(\theta)}{a} \frac{\eta_l \eta_e}{\eta_l + \eta_e},$$

⁵For a sensitivity analysis as to how the optimal asymptotic tax rate varies with different parameter, see Diamond (1998).

which means that now optimal taxes will converge to approximately 0.33.

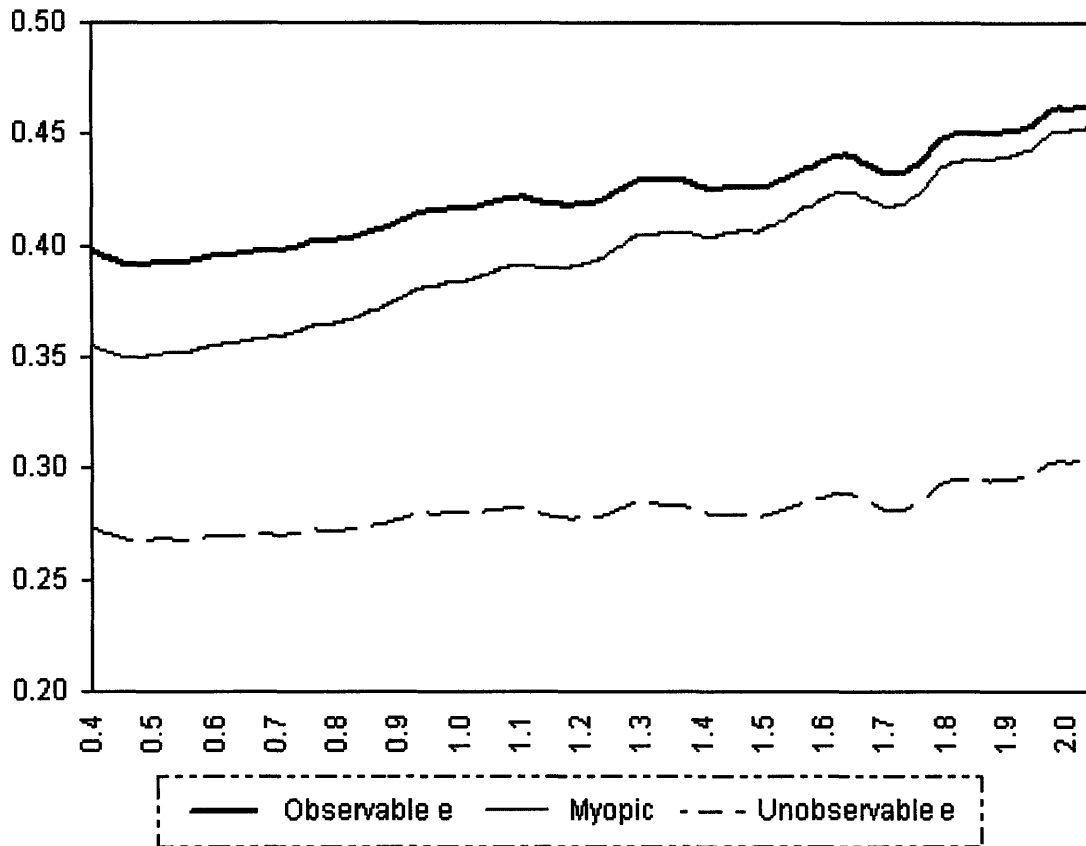


Figure 2-3: Estimated Optimal Tax Schedule

Figure 2-3 plots the simulated optimal tax schedules under the three different assumptions for education. The optimal tax for very low values of ability is not reported, since the estimate for these values is probably less accurate since observations with very low wages were dropped and did not consider people who chose not to work were not considered. Such a detailed analysis is beyond the scope of this paper. For very high levels of ability, the tax schedule will continue to rise until we reach the limit tax rates calculated above. However, the estimated series gets significantly noisier since the number of observations with very high ability is very low.

As mentioned before, the existence of unobservable education will tend to make

tax rates lower, since there are additional distortions arising from taxation⁶. This is indeed the case in our simulation. In the graph below we can see that taxes are at their lowest when education is unobservable. When education is observable, taxes are at their highest, but the choice of education is not affected by them since education is being subsidized at the same rate, which means that even though taxes are very high, they are not as distortionary as they would be if education was not observable.

Finally, in the case where the planner ignores the effect of taxes on education, taxes are higher than when unobservable education is taken into account exactly because this distortion is not considered and lower than when education is observable because individuals did not invest as much in education which means that society will be poorer, which decreases the $B(\theta)$ component of the tax schedule.

At this point it is useful to take stock and compare our results to those of Kapika (2004). He finds that taking the distortion of taxes on education into account should significantly reduce taxes in the US. Our findings indeed support this conclusion if we believe that we are switching from a tax system that totally ignores taxation effects on education to one that takes them into account, but cannot observe education choices. However, if we can switch to a system where education is observable, then income taxes should instead go up and additional subsidies to education should be provided.

2.3.4 Actual Taxes in the US

Using the same CPS data as before, the expected wage and the expected tax rate⁷ were computed for any given individual given his schooling level. This gives us an idea of what the optimal marginal subsidy rate should be for each level of schooling under the monetary cost of education model.

It is important to emphasize again that the model does not take into consideration any redistribution from children of rich parents to children of poor parents. The model

⁶This need not be the case in a more general model. Diamond and Mirrlees (2002) provide an example where adding an additional distortive decision variable need not reduce optimal taxes if income effects are strong enough. Even in this quasi-linear model it is not clear that with unobservable education taxes will be lower for all types.

⁷These tax rates refer to single individuals.

	Expected Wages	Marginal Taxes
No School	17,644	15%
Grade 1 to 4	20,479	17%
Grade 5 to 8	21,970	17%
Grade 9	23,014	17%
Grade 10	24,371	17%
Grade 11	24,184	17%
Grade 12, no diploma	26,084	18%
High school diploma	32,143	19%
Some college	37,336	21%
Associate degree, occupational program	38,509	21%
Associate degree, academic program	41,386	22%
Bachelor's degree	57,555	24%
Master's degree	67,265	25%
Professional degree	131,103	28%
Doctorate degree	102,958	27%

Figure 2-4: Expected Wages and Taxes for each Schooling Level

implicitly assumes zero initial endowment for everyone and perfect credit markets. It does not take intergenerational redistribution issues into consideration. Also, this analysis focuses only on the level of education achieved, and not on its quality. Finally, if income tax rates are not set optimally, then in general it will not be optimal to set marginal subsidies on education equal to marginal income taxes. For these reasons, these numbers should be taken as the direct application of a stylized model and not as a policy recommendation.

Another interesting feature of the US tax system is that the marginal tax rate is increasing (displaying a similar pattern to the one simulated in Figure 3) and reaches its maximum at 35%. Given that the distribution of skills seems to follow a Pareto distribution with parameter $a = 3$, we can now find the elasticity of labor substitution for which this asymptotic tax is optimal according to the following expression

$$\eta_l B(\theta) = \frac{T_w}{1 - T_w} a \simeq 1.62,$$

which for our usual assumption that G' converges to zero gives us an implied elasticity of labor supply of 1.63, which seems to be too high. To get a more reasonable value of

around 0.5 we would need $B(\theta)$ to be almost 0.5, which means that the government cares a lot about high earners. Another possibility is that education is actually not totally observable.

2.4 Concluding Remarks

The model proposed in this chapter suggests that if education is observable, then its monetary costs should be subsidized at the same marginal rate as we are taxing the income of individuals in each ability level. Non-monetary costs should also be subsidized, but it is not clear exactly to what extent.

For these results to hold it is crucial that the cost of education is observable (or at least mostly so) and that ability and education are complements, although the particular functional form for the production function may vary.

Although full flexibility of the tax function is allowed, the results seem to suggest that there is a relatively simple way to operationalize them (at least regarding monetary costs), which is to allow educational expenses to be deducted from future income (for example, by allowing tax deductible student loan payments). Current US income taxes allow tuition and fees to be deducted from current income (or from parents' income) and allow student loan interest deduction. Although this measure is clearly a subsidy on education, it is not clear that its form is the best. For example, if a self financed student with no income at time of education gets funding through a student loan, the subsidy he receives is probably suboptimal, since he only gets deduction on his interest payments. On the other hand, children from parents with a high income will receive a high subsidy on education even if they end up paying a low marginal tax due to lower ability. However, a more complete discussion of these issues would have to incorporate the effects of inherited wealth, which are outside the scope of this paper.

Another issue is that the people at the bottom of the distribution are not the ones receiving the maximum marginal subsidy for education. As in the original Mirrlees model, marginal tax rates at the bottom of the distribution are zero, which means

that the marginal subsidy will also be quite low for the lowest types. This seems to be in strong conflict with the common view that people should have the same opportunities and that education is crucial in providing that. One answer might be that in this model being at the bottom of the distribution has nothing to do with low initial wealth that might prevent people from getting education, but is rather related to having such low ability that investment in education has very low return. Also, our results do not mean that everyone should not be getting a minimum level of education. How this education should be financed is indeterminate, though. The lowest type will in general receive a negative tax, but it can be paid either as an income transfer or free education up to a certain level.

2.5 Appendix

2.5.1 Derivation of the Optimal Tax First-Order Conditions

Assuming that the first order conditions for the consumer's problem fully characterize their optimal choice, we can write the government's problem as

$$\begin{aligned}
 \max \quad & \int_{\underline{\theta}}^{\bar{\theta}} G(U(\theta e_{\theta} l_{\theta} - T(\theta e_{\theta} l_{\theta}, e_{\theta}), e_{\theta}, l_{\theta})) f(\theta) d\theta \\
 \text{subject to} \quad & E = \int_{\underline{\theta}}^{\bar{\theta}} T(\theta e_{\theta} l_{\theta}, e_{\theta}) f(\theta) d\theta \\
 & U_e = -U_y [\theta l [1 - T_w(\theta e y, e)] - T_e(\theta e y, e)] \\
 & U_l = -U_y \theta e [1 - T_w(\theta e y, e)].
 \end{aligned}$$

We will now make the following change of variables

$$t(\theta) = T(\theta e_{\theta} l_{\theta}, e_{\theta}) \text{ with}$$

$$u(\theta) = U(\theta e_{\theta} l_{\theta} - t(\theta), e_{\theta}, l_{\theta}) \text{ and } u'(\theta) = U_y e_{\theta} l_{\theta} [1 - T_w] = -\frac{U_l l_{\theta}}{\theta},$$

which means we can write the government's problem as

$$\begin{aligned}
& \max \int_{\underline{\theta}}^{\bar{\theta}} G(u(\theta)) f(\theta) d\theta \\
\text{subject to } & E = \int_{\underline{\theta}}^{\bar{\theta}} t(\theta) f(\theta) d\theta \\
& u(\theta) = U(\theta e_{\theta} l_{\theta} - t(\theta), e_{\theta}, l_{\theta}) \\
& -u'(\theta) = U_l(\theta e_{\theta} l_{\theta} - t(\theta), e_{\theta}, l_{\theta}) l_{\theta} / \theta.
\end{aligned}$$

The Hamiltonian for this problem is

$$\begin{aligned}
H = & f(\theta) \{ G(u(\theta)) + \rho t(\theta) \} + \frac{\sigma_{\theta} l_{\theta}}{\theta} U_l(\theta e_{\theta} l_{\theta} - t(\theta), e_{\theta}, l_{\theta}) \\
& - \mu_{\theta} \{ u(\theta) - U(\theta e_{\theta} l_{\theta} - t(\theta), e_{\theta}, l_{\theta}) \}
\end{aligned}$$

and the first order conditions are given by

$$\begin{aligned}
\sigma'_{\theta} &= \mu_{\theta} - f(\theta) G'(u(\theta)) \\
\rho f(\theta) &= \mu_{\theta} U_y + \frac{\sigma_{\theta} l_{\theta}}{\theta} U_{ly} \\
\mu_{\theta} [\theta e_{\theta} U_y + U_l] &= -\frac{\sigma_{\theta} l_{\theta}}{\theta} \left[\frac{U_l}{l_{\theta}} + U_{ll} + \theta e_{\theta} U_{ly} \right] \\
\mu_{\theta} [\theta l_{\theta} U_y + U_e] &= -\frac{\sigma_{\theta} l_{\theta}}{\theta} [U_{le} + \theta l_{\theta} U_{ly}].
\end{aligned}$$

Plugging in the consumer's first order conditions, we can rewrite these expressions as

$$\begin{aligned}
\frac{T_w}{1 - T_w} &= \frac{\sigma_{\theta}}{\theta \mu_{\theta}} A, \text{ where } A = 1 + l_{\theta} \frac{U_{ll}}{U_l} + \theta e_{\theta} l_{\theta} \frac{U_{ly}}{U_l} \\
\frac{T_w}{1 - T_w} - \frac{T_e U_y}{U_e (1 - T_w)} &= \frac{\sigma_{\theta}}{\theta \mu_{\theta}} B, \text{ where } B = l_{\theta} \frac{U_{le}}{U_e} + \theta l_{\theta}^2 \frac{U_{ly}}{U_e}.
\end{aligned}$$

Using these expressions we can solve for T_e as a function of T_w

$$T_e = T_w \frac{U_e}{U_y} \frac{A - B}{A}.$$

Chapter 3

Taxation of Entrepreneurial Capital in a Deterministic Setting

This chapter considers a Ramsey model of linear taxation for an economy with capital and two kinds of labor. If the government can only observe the joint return of capital and entrepreneurial labor, then there will be positive capital income taxation, even in the long run. This happens because the only way to tax entrepreneurial labor is by also taxing capital. An example shows conditions under which the labor tax is higher than the capital tax.

3.1 Introduction

In a Ramsey economy with capital and one type of labor, Chamley (1986) shows that it is optimal not to use capital taxation in the long run, so that only labor taxes are used to finance government spending. The reason for this is that since labor will bear the full incidence of both taxes, it is more efficient not to use capital taxation, which also distorts the intertemporal wedge, as well as the intratemporal decision between consumption and leisure.

An interesting departure from this framework is the case where there are two types of labor in an economy. Optimal capital taxes will in general not be zero in the long run if one labor input is not taxable (Correia 1996) or both kinds of labor have to

be taxed at the same rate (Jones, Manuelli, and Rossi 1997). However, the sign of the optimal capital tax remains ambiguous and strongly depends on functional form assumptions.

This chapter also considers a model with two types of labor, but assumes that one kind of labor, which can be interpreted as entrepreneurial labor, has to be taxed at the same rate as capital. Under these conditions, the optimal plan will have positive capital taxes in the long run. The intuition for this result is that the government would like to tax both kinds of labor and not tax capital in the long run. However, since the only way to tax entrepreneurial labor is by also taxing capital, the government will choose to use capital taxation, so that it can tax entrepreneurial labor.

This deterministic model allows us to capture one reason why the existence of unobservable entrepreneurial labor may lead to positive capital taxation in the long run. A more complete study of entrepreneurial investment will also need to take into account the uncertainty that characterizes these activities, and possible insurance mechanisms to overcome it. For a reference on these issues, see Albanesi (2006), who finds a different mechanism for capital taxation, which might be positive or negative, depending on whether higher capital tightens or loosens the incentive compatibility for the entrepreneur's effort.

The chapter proceeds as follows. Section 3.2 sets up the model and describes the optimality conditions for agents in the economy. Section 3.3 derives the optimal plan for the benevolent government, which entails positive capital taxation in the long run. Section 3.4 shows an example with sufficient conditions for capital taxes to be positive, but lower than labor taxes. Section 3.5 concludes. The appendix in section 3.6 derives a baseline model where both kinds of labor income can be taxed at different rates, as well as some extensions of the model with one type of labor.

3.2 Model Setup

When entrepreneurs work in their own business, it is hard to disentangle which part of the firm's profits is due to their effort, and which is a pure return on capital. Thus,

the government may not be able to charge different tax rates for the return on capital and entrepreneurial labor.

This section sets up a model of linear Ramsey taxation when there is one capital good and two labor inputs. The return from one of the labor inputs cannot be observed separately from the return on capital, which means that they are subject to the same tax rate. This labor input can be interpreted as entrepreneurial labor. The other labor input is fully observable and is subject to a labor tax. The government is benevolent and needs to collect revenue to finance public spending of g in every period.

3.2.1 Households

There is a continuum of measure one of identical households who derive utility from consumption c_t , observable labor l_t , and entrepreneurial labor n_t . They discount the future at rate β , so that each household's lifetime utility is given by

$$\sum_{t=0}^{\infty} \beta^t u(c_t, l_t, n_t).$$

Assume for simplicity that the utility function takes the following form

$$u(c, l, n) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{l^{1+\alpha}}{1+\alpha} - \frac{n^{1+\eta}}{1+\eta} \text{ with } \sigma, \alpha, \eta > 0.$$

Households can spend their income in consumption goods, or they can save using capital k_t or government bonds b_t . Let q_t be the price of a bond that pays one unit of consumption good tomorrow. Households receive income from their labor and from their investment in capital and bonds. Let $(1 - \tau_t^l)w_t^l$ be the after tax return on observable labor. Entrepreneurial labor income w_t^n and capital income r_t are subject to the same tax rate τ_t^k . Given this, the households face the following budget constraint every period

$$c_t + k_{t+1} + q_t b_{t+1} = (1 - \tau_t^l)w_t^l l_t + (1 - \tau_t^k)(w_t^n n_t + r_t k_t) + b_t,$$

as well as a no-ponzi condition that guarantees that households cannot roll over their debt forever.

Households maximize their utility subject to their budget constraint. The solution to this problem is given by the first order conditions for consumer optimality.

The intertemporal conditions for consumption and capital accumulation are

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = \frac{1}{q_t} = r_{t+1}(1 - \tau_{t+1}^k).$$

The intratemporal conditions for consumption and the two kinds of labor are

$$-\frac{u_{n,t}}{u_{c,t}} = w_t^n(1 - \tau_t^k) \text{ and } -\frac{u_{l,t}}{u_{c,t}} = w_t^l(1 - \tau_t^l).$$

Together with the budget constraint, these conditions fully characterize the solution to the households' problem.

Notice that the same tax rate τ_t^k that determines the distortion in the entrepreneurial labor choice also controls the wedge in capital accumulation. As we will see in the next section, this will imply that the government cannot choose different distortions for these two margins, which will lead to positive capital taxation in the long run.

3.2.2 Firms and Market Clearing

Each period firms maximize profits given the before taxes prices for capital r_t and the labor inputs w_t^l and w_t^n . They have access to the production function $F(k_t, l_t, n_t)$, which has constant returns to scale and is concave. Furthermore, assume $F_{kx} \geq 0$ for $x = n, l$, so that increasing capital never decreases the productivity of either type of labor. Market clearing implies that factor prices must equal the marginal productivity of each factor

$$r_t = F_{k,t} \quad w_t^l = F_{l,t} \quad w_t^n = F_{n,t}.$$

The resource constraint for the economy must also be met every period

$$c_t + g + k_{t+1} = F(k_t, l_t, n_t).$$

3.3 Optimal Plan

Given the households' budget constraint and the resource constraint, market clearing conditions ensure that the government's budget constraint is always met, which means we do not need to take it explicitly into account. Furthermore, following Lucas and Stokey (1983), we can simplify the households' budget constraint using the optimality conditions for households and firms, and rewrite the budget constraint as the following implementability condition

$$\sum_{t=0}^{\infty} \beta^t m(c_t, n_t, l_t) = u_c(c_0, n_0, l_0) [F_k(k_0, n_0, l_0)(1 - \tau_0^k)k_0 + b_0] \equiv W_0$$

with $m(c, n, l) = u_c c + u_n n + u_l l$.

In general, if an allocation meets the implementability condition and the resource constraint, it can be implemented as a competitive equilibrium for certain prices and taxes. However, for this problem, since the tax on entrepreneurial labor is the same as the tax on capital, we also need to impose the following observability condition to make sure that the same capital tax satisfies the households' optimality conditions for capital and entrepreneurial labor

$$u_{n,t+1} \beta F_{k,t+1} + u_{c,t} F_{n,t+1} = 0.$$

For the initial period, the observability condition is given by

$$M_0 \equiv u_{n,0} + u_{c,0} F_{n,0} (1 - \tau_0^k) = 0.$$

The benevolent planner thus maximizes the welfare of a representative consumer subject to the implementability condition (associated to multiplier μ), the resource constraint for every period (associated with multiplier $\beta^t \rho_t$), and observability condi-

tion (associated with multiplier $\beta^t \gamma_{t+1}$). The Lagrangean for this problem is

$$L = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t, l_t) - \gamma_{t+1} [u_{n,t+1} \beta F_{k,t+1} + u_{c,t} F_{n,t+1}] + \\ \mu m(c_t, n_t, l_t) - \rho_t [c_t + g_t + k_{t+1} - F(k_t, n_t, l_t)] \end{array} \right\} - \mu W_0 - \frac{\gamma_0 M_0}{\beta}.$$

The first order condition for consumption for $t \geq 1$ is given by

$$u_{c,t} + \mu m_{c,t} - \rho_t = \gamma_{t+1} u_{cc,t} F_{n,t+1}.$$

The first order condition for entrepreneurial labor for $t \geq 1$ is

$$u_{n,t} + \mu m_{n,t} + \rho_t F_{n,t} = \frac{\gamma_t}{\beta} (u_{nn,t} \beta F_{k,t} + u_{n,t} \beta F_{kn,t} + u_{c,t-1} F_{nn,t}).$$

And the first order condition for capital for $t \geq 1$ is

$$\rho_t - \rho_{t+1} \beta F_{k,t+1} = -\gamma_{t+1} (u_{n,t+1} \beta F_{kk,t+1} + u_{c,t} F_{nk,t+1}).$$

We can now show that in steady state there will be positive capital taxation. To see this, we will start by showing that, in any steady state, γ_t cannot be zero, and both γ_t and ρ_t must be constant. Under these conditions, the first order condition for capital accumulation implies that there must be positive capital taxation in the long run.

To see that in steady state, γ_t cannot be zero, assume there is a steady state with $\gamma_t = 0$. Then the planner's first order condition for consumption implies that ρ_t must be constant since $\rho_t = \rho = u_c + \mu m_c$. Given this, the first order condition for capital implies that $\beta F_k = 1$. If this is the case, then the observability condition can only be met if

$$F_n = -\frac{u_n}{u_c}.$$

On the other hand, with $\gamma_t = 0$, the planner's first order conditions for consump-

tion and entrepreneurial labor imply that

$$F_n = -\frac{u_n + \mu m_n}{u_c + \mu m_c} = -\frac{u_n}{u_c} \frac{1 + \mu(1 - \sigma)}{1 + \mu(1 + \eta)} < -\frac{u_n}{u_c}.$$

Thus γ_t cannot be zero in steady state since this would entail taxing capital at a different rate from entrepreneurial labor, which is not allowed in this model.

Let us now check that, in steady state, γ_t and ρ_t must be constant. To see this, solve for ρ_t and ρ_{t+1} as a function of γ_{t+1} using first order conditions for consumption and labor in the following way

$$\begin{aligned}\rho_t(\gamma_{t+1}) &= u_c + \mu m_c - \gamma_{t+1} u_{cc} F_n \\ \rho_{t+1}(\gamma_{t+1}) &= [\gamma_{t+1}(u_{nn}\beta F_k + u_n\beta F_{kn} + u_c F_{nn})/\beta - u_n - \mu m_n]/F_n.\end{aligned}$$

Plug into first order condition for capital

$$\rho_t(\gamma_{t+1}) - \rho_{t+1}(\gamma_{t+1})\beta F_k = -\gamma_{t+1}(u_n\beta F_{kk} + u_c F_{nk}).$$

We will get an equation that is linear in γ_{t+1} and thus uniquely defines γ_{t+1} , which implies that in steady state both γ_t and ρ_t must be constant.

Since γ_t must be constant and positive in steady state, the planner's first order condition for capital together with the households' first order condition for capital implies that there will be a positive capital tax as long as $u_n\beta F_{kk} + u_c F_{nk} > 0$. Using the observability condition, this condition can be rewritten as

$$u_c F_n \left[\frac{F_{nk}}{F_n} - \frac{F_{kk}}{F_k} \right] > 0,$$

which, given our assumption for the production function, will always be met.

3.4 An Example

In the previous section we have seen that if entrepreneurial labor must be taxed at the same rate as capital, then it is optimal to have positive capital taxation in the long run, so that entrepreneurial capital can bear some of the tax burden. However, using this kind of taxation leads to an undesirable distortion of the intertemporal wedge. Thus, although capital taxes are positive, they will tend to be smaller than if they were a pure tax on entrepreneurial labor.

The example that follows shows conditions under which the tax on both kinds of labor income would be the same if they were both observed separately from capital income, but if entrepreneurial labor income is observed jointly with capital income, then it will be taxed at a positive rate, but lower than that on observable labor. The reason for this is that since the capital tax is creating an additional distortion on capital accumulation, then it will be optimal for it to be smaller than if it was a pure tax on entrepreneurial labor.

Assume that $\alpha = \eta$, so that both kinds of labor have the same elasticity of substitution. Assume also that the production function is given by

$$F(k, l, n) = k^\phi(l + n)^{1-\phi} \text{ with } 0 < \phi < 1.$$

Take the planner's first order condition for entrepreneurial labor and divide it by $u_{n,t}$ to get the following expression

$$1 + \mu(1 + \eta) = \gamma_t \frac{u_{nn,t}}{u_{n,t}} F_{k,t} + \frac{A^n}{u_{n,t}},$$

where $A^n \equiv u_{n,t} \gamma_t F_{kn,t} + \gamma_t u_{c,t-1} F_{nn,t} / \beta - \rho_t F_{n,t}$.

For observable labor, the first order condition can be written as

$$1 + \mu(1 + \eta) = \frac{A^l}{u_{l,t}},$$

where $A^l \equiv u_{n,t}\gamma_t F_{kl,t} + \gamma_t u_{c,t-1} F_{nl,t} / \beta - \rho_t F_{l,t}$.

Given the assumption that the two types of labor are perfect substitutes in production, it follows that $A^n = A^l = A$. Thus, it must be the case $u_{n,t} < u_{l,t}$ since

$$u_{n,t} = \frac{A + \gamma_t u_{nn,t} F_{k,t}}{1 + \mu(1 + \eta)} < \frac{A}{1 + \mu(1 + \eta)} = u_{l,t}.$$

Since $w_t^n = w_t^l = w_t$, the households's optimality conditions for labor imply that the capital tax is always lower than the labor tax

$$\begin{aligned} \frac{u_{n,t}}{1 - \tau_t^k} &= -w_t u_{c,t} = \frac{u_{l,t}}{1 - \tau_t^l}. \\ \frac{1 - \tau_t^k}{1 - \tau_t^l} &= \frac{u_{n,t}}{u_{l,t}} > 1 \text{ since } u_{n,t} < u_{l,t} < 0. \end{aligned}$$

The reason for this is that in this case the government would like to set the same tax for both kinds of labor. However, since taxing entrepreneurial labor also taxes capital, it is optimal to reduce this kind of taxation relative to pure labor taxation.

If the government would like to initially tax entrepreneurial labor at different rates, then the optimal capital tax will be higher than the labor tax if the following condition is met

$$u_{n,t} = \frac{A - \gamma_t u_{nn,t} F_{k,t}}{1 + \mu(1 + \eta)} > \frac{A}{1 + \mu(1 + \alpha)} = u_{l,t}.$$

3.5 Concluding Remarks

In general, if labor and capital taxes are available, the government will use only labor taxation in the long run. The reason for this is that the full incidence of the capital tax is on labor since the intertemporal rate of return is pinned down by the discount rate. Furthermore, capital taxes also distort the intertemporal margin, which means that labor taxes are the most efficient way to collect revenue. However, if there are two types of labor, one of which must face the same tax as capital, then the only way

to tax it is by using capital taxation. In this case, the incidence of the capital tax is no longer the same as that of the pure labor tax, which means that it is optimal to use both kinds of taxation.

3.6 Appendix

3.6.1 Optimal Taxation with Two Taxable Labor Inputs

If different taxes are allowed for capital income and entrepreneurial labor income, then the optimal plan no longer needs to meet the constraint

$$u_{n,t+1}\beta F_{k,t+1} + u_{c,t}F_{n,t+1} = 0.$$

The new Lagrangean for this problem is

$$L = \sum_{t=0}^{\infty} \beta^t \{u(c_t, n_t, l_t) + \mu m(c_t, n_t, l_t) - \rho_t [c_t + g_t + k_{t+1} - F(k_t, n_t, l_t)]\} - \mu W_0.$$

This means that the optimality conditions will be the same as before, but with the multiplier γ set equal to zero, so that for our functional form, the optimal tax on capital is zero and taxes for both kinds of labor are given by

$$\begin{aligned} 1 + \tau^n &= \frac{1 + \mu(1 - \sigma)}{1 + \mu(1 + \eta)} \\ 1 + \tau^l &= \frac{1 + \mu(1 - \sigma)}{1 + \mu(1 + \alpha)}. \end{aligned}$$

Thus, the Chamley (1986) holds with two kinds of labor as long as they can be taxed separately.

3.6.2 Optimal Taxation with One Untaxable Labor Input

The Chamley (1986) model had one labor input n_t that was subject to a linear tax. Consider instead that labor cannot be taxed.

The utility function is now given by

$$u(c, n) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{n^{1+\eta}}{1+\eta} \text{ with } \sigma, \eta > 0.$$

The production function $F(k, n)$ is increasing, concave, and exhibits constant returns to scale in capital and labor.

The implementability condition can be derived as in the model with two labor inputs and is given by

$$\sum_{t=0}^{\infty} \beta^t m(c_t, l_t) = W_0.$$

If the labor input cannot be taxed, then the following observability constraint needs to be met

$$u_{n,t} + w_t u_{c,t} = 0.$$

The planner maximizes utility subject to the implementability condition, the observability constraint, and the new resource constraint, so that Lagrangean for the planner's problem is given by

$$L = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t) + \mu m(c_t, n_t) - \gamma_t [u_{n,t} + F_{n,t} u_{c,t}] \\ -\rho_t [c_t + g_t + k_{t+1} - F(k_t, n_t)] \end{array} \right\} - \mu W_0.$$

The first order condition for consumption for $t \geq 1$ is given by

$$u_{c,t} + \mu m_{c,t} - \rho_t = \gamma_t F_{n,t} u_{cc,t}.$$

The first order condition for labor for $t \geq 1$ is

$$u_{n,t} + \mu m_{n,t} + \rho_t F_{n,t} = \gamma_t (u_{nn,t} + F_{nn,t} u_{c,t}).$$

And the first order condition for capital for $t \geq 1$ is

$$\rho_t - \rho_{t+1} \beta F_{k,t+1} = -\gamma_{t+1} F_{nk,t+1} u_{c,t+1}.$$

As before, in steady state, the multiplier γ has to be positive, otherwise there would be positive capital taxation, which is not allowed in this model. Thus, the capital tax cannot be zero. Furthermore, since $F_{nk} > 0^1$, the capital tax must be strictly positive and will be given by

$$1 - \tau^k = \frac{1}{\beta F_k} \text{ with } \beta F_k = 1 + \frac{\gamma}{\rho} F_{nk} u_c.$$

Notice that this result is very similar to that of Correia (1996), where there was one untaxable labor input. Here, the untaxable labor input is the only one, which means that its cross derivative with capital in the production function is unambiguously signed, which leads to the conclusion that there must be positive capital taxation in the long run.

3.6.3 Optimal Joint Capital and Labor Tax

Assume now that there is one labor input that must be taxed at the same rate as capital. Then, as in the model with entrepreneurial labor, the following condition must be imposed

$$u_{n,t+1} \beta F_{k,t+1} + u_{c,t} F_{n,t+1} = 0.$$

The Lagrangean for this problem is

$$L = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t) - \gamma_{t+1} [u_{n,t+1} \beta F_{k,t+1} + u_{c,t} F_{n,t+1}] + \\ \mu m(c_t, n_t) - \rho_t [c_t + g_t + k_{t+1} - F(k_t, n_t)] \end{array} \right\} - \mu W_0 - \frac{\gamma_0 M_0}{\beta}.$$

The first order condition for consumption for $t \geq 1$ is given by

$$u_{c,t} + \mu m_{c,t} - \rho_t = \gamma_{t+1} u_{cc,t} F_{n,t+1}.$$

¹Note that now that there are only two production factors we do not need to assume this. Since the production function is concave and has constant returns to scale, it must be the case that $F_{nk} > 0$.

The first order condition for labor for $t \geq 1$ is

$$u_{n,t} + \mu m_{n,t} + \rho_t F_{n,t} = \frac{\gamma_t}{\beta} (u_{nn,t} \beta F_{k,t} + u_{n,t} \beta F_{kn,t} + u_{c,t-1} F_{nn,t}).$$

And the first order condition for capital for $t \geq 1$ is

$$\rho_t - \rho_{t+1} \beta F_{k,t+1} = -\gamma_{t+1} (u_{n,t+1} \beta F_{kk,t+1} + u_{c,t} F_{nk,t+1}).$$

As in the model with entrepreneurial capital, the multiplier γ cannot be zero in the long run, which means that if there is a unique income tax rate, then capital will be taxed in the long run.

Chapter 4

Optimal Taxation with Unobservable Investment in Human Capital

This chapter considers a Ramsey model of linear taxation where the government cannot distinguish between human capital and labor income. In this context, if individual investment in human capital is unobservable, then it is optimal to tax human capital at a positive rate, even in the long run. This is true even if consumption taxes are available. Whether physical capital should be taxed depends on its degree of complementarity with human capital versus raw labor.

4.1 Introduction

In a Ramsey linear taxation model, if the government is allowed to tax both labor and capital income, then in the long run it is optimal to converge to zero taxes on capital income and collect all revenue through labor taxes (Chamley 1986). However, in reality, labor income also includes a large component of human capital, which means that human capital will be taxed in the long run if labor taxes are positive.

Jones, Manuelli, and Rossi (1997) and Judd (1999) show that if the government can distinguish between pure consumption and human capital investment, then it

can use this information to offset the distortion that labor taxation causes on human capital accumulation. Under this condition, it is optimal not to distort either human or physical capital accumulation in the long run.

This chapter assumes that the government cannot distinguish between final consumption and expenditures on human capital. If this is the case, then human capital will in general be taxed in the long run, even if consumption taxes are available. Physical capital may be taxed or subsidized, depending on its degree of complementarity with human capital relative to raw labor.

The chapter proceeds as follows. Section 4.1 presents a model of optimal taxation in an economy with human capital and endogenous growth, which is a simplified version of the model in Jones, Manuelli, and Rossi (1997). Section 4.2 introduces the assumption that human capital expenditures cannot be observed separately from final consumption and derives the optimal tax scheme under these conditions. Section 4.3 concludes. The appendix in section 4.4 introduces an economy without endogenous growth, and finds that the previous results are not sensitive to this assumption.

4.2 Benchmark Model

This section presents a model of human capital taxation which is a simplified version of the model in Jones, Manuelli and Rossi (1997). The government can use capital and labor income linear taxation τ^k and τ^l , as well as differentiated consumption taxes on final consumption τ^c and human capital expenditures τ^h to finance a given amount of spending in the public good g every period.

4.2.1 Households

There is a continuum of measure one of identical households who derive utility from consumption c_t and labor n_t . Per period utility $u(c_t, n_t)$ is increasing in consumption and decreasing in labor and globally concave. Households discount the future at rate

β , so that their lifetime utility is given by

$$\sum_{t=0}^{\infty} \beta^t u(c_t, n_t).$$

Households have to meet their intertemporal budget constraint

$$\sum_{t=0}^{\infty} p_t [c_t(1 + \tau_t^c) + k_{t+1} + h_{t+1}(1 + \tau_t^h) - r_t(1 - \tau_t^k)k_t - w_t(1 - \tau_t^l)l_t] = p_0 b_0.$$

In this very stylized economy human capital accumulation is done simply by consuming some amount of final goods, and there is no labor component in human capital, as well as no depreciation. Depreciation could be added without changing the nature of our results, so this is done for notational simplicity. Assuming that there is no labor cost allows us to focus on the pure investment component of human capital, given that the results in Jones, Manuelli and Rossi (1997) hold regardless of whether labor is needed to produce human capital. Likewise, physical capital only requires investment of final goods.

Households spend their income in consumption goods, as well as in physical and human capital for the next period k_{t+1} and h_{t+1} . Households receive income from capital and effective labor l_t , which results from the combination of human capital and raw labor according to $l_t = h_t n_t$. The initial level of government debt that is held by households is given by b_0 and the (before taxes) price of consumption in period t is p_t .

Households maximize their utility subject to their budget constraint. Given that this problem is concave, the first order conditions are necessary and sufficient for optimality of the households' choice.

The optimality condition for physical capital accumulation is given by

$$r_{t+1}(1 - \tau_{t+1}^k) = \frac{p_t}{p_{t+1}} = \frac{u_{c,t}}{\beta u_{c,t+1}} \frac{1 + \tau_{t+1}^c}{1 + \tau_t^c}.$$

The optimality condition for human capital accumulation is given by

$$\frac{p_t}{p_{t+1}} = w_{t+1}n_{t+1} \frac{1 - \tau_{t+1}^l}{1 + \tau_t^h}.$$

Finally, the optimality condition for the labor and consumption choice is

$$u_{n,t} = -u_{c,t}w_t h_t \frac{1 - \tau_t^l}{1 + \tau_t^c}.$$

Together with the budget constraint, these conditions fully characterize the households' behavior.

4.2.2 Firms and Market Clearing

Each period firms maximize profits given the (before taxes) prices for labor w_t and capital r_t . They have access to the production function $F(k_t, l_t)$, which has constant returns to scale and decreasing marginal productivity of capital and effective labor. Market clearing implies that factor prices must equal the marginal productivity of each factor

$$w_t = F_l(k_t, l_t)$$

$$r_t = F_k(k_t, l_t).$$

The resource constraint for the economy must also be met every period

$$c_t + g + k_{t+1} + h_{t+1} = F(k_t, l_t).$$

4.2.3 Optimal Plan

Given the households' budget constraint and the resource constraint, market clearing conditions ensure that the government's budget constraint is always met, which means we do not need to take it explicitly into account. Furthermore, following Lucas and Stokey (1983), we can simplify the households' budget constraint using the optimality

conditions for households and firms, and write it as the following implementability condition

$$\sum_{t=0}^{\infty} \beta^t c_t u_{c,t} = \frac{u_{c,0}}{1 + \tau_0^c} [F_{k,0}(1 - \tau_0^k)k_0 + F_{l,0}(1 - \tau_0^l)l_0 + b_0] \equiv W_0.$$

To see this, note that using the first order conditions for household optimality we can show that the following equalities must hold

$$\begin{aligned} \sum_{t=0}^{\infty} p_t [k_{t+1} - r_t(1 - \tau_t^k)k_t] &= p_0 r_0 (1 - \tau_0^k) k_0 \\ \sum_{t=0}^{\infty} p_t [h_{t+1}(1 + \tau_t^h) - w_t(1 - \tau_t^l)l_t] &= p_0 w_0 (1 - \tau_0^l) l_0 \\ \sum_{t=0}^{\infty} p_t [c_t(1 + \tau_t^c)] &= p_0 \frac{1 + \tau_0^c}{u_{c,0}} \sum_{t=0}^{\infty} \beta^t c_t u_{c,t}. \end{aligned}$$

Given this, it is straightforward to see that the households' budget constraint is equivalent to the implementability condition.

Any allocation that meets both the implementability condition and the resource constraint can be implemented as a competitive equilibrium in this economy since there exist prices and taxes that guarantee that the optimality conditions for the households and firms are met. Thus, a benevolent government will maximize social welfare subject to those two constraints

$$\begin{aligned} \max \quad & \sum_{t=0}^{\infty} \beta^t u(c_t, n_t) \\ \text{subject to} \quad & c_t + g + k_{t+1} + h_{t+1} = F(k_t, l_t) \\ & \sum_{t=0}^{\infty} \beta^t c_t u_{c,t} = W_0. \end{aligned}$$

Let μ be the multiplier on the implementability condition and let $\beta^t \rho_t$ be the multiplier on the resource constraint in period t . Then the optimality conditions for

c_t , n_t , k_t , and h_t , for $t \geq 1$, are given by

$$\begin{aligned} u_{c,t} + \mu(u_{c,t} + c_t u_{cc,t}) &= \rho_t \\ u_{n,t} + \mu c_t u_{cn,t} &= -\rho_t F_{l,t} h_t \\ \rho_{t+1} \beta F_{k,t+1} &= \rho_t \\ \rho_{t+1} \beta F_{l,t+1} n_t &= \rho_t. \end{aligned}$$

In steady state ρ will be constant, which means that neither physical nor human capital accumulation will be distorted. However, the intratemporal wedge between labor and consumption will in general be distorted as we can see in the expression below

$$\frac{1 - \tau^n}{1 + \tau^c} = -\frac{u_n}{u_c w h} = \frac{1 + \mu(1 + c u_{cc}/u_c)}{1 + \mu c u_{cn}/u_n}.$$

Notice that there are multiple taxes that implement this allocation, since there is one additional degree of freedom when we are choosing the taxes that meet the optimality conditions. As Jones, Manuelli and Rossi (1997) point out, one possibility is to have zero capital and labor taxes, as well as zero tax on human capital investments. Consumption taxes may be zero for specific functional forms, but will be (strictly) positive as long as utility is separable between labor and consumption and the coefficient of risk aversion is (strictly) higher than one.

Another possible tax scheme that implements the optimal solution to this problem would be to have zero consumption and capital taxes, together with a positive labor tax and a subsidy on human capital expenditures.

From now on, assume that the utility function takes the following form

$$u(c, n) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{n^{1+\gamma}}{1+\gamma} \text{ with } \sigma > 1.$$

4.3 Unobservable Investment in Human Capital

To implement the optimal solution with observable investment in human capital, we had one degree of freedom in choosing taxes, which translates into the fact that

any of the available tax rates τ_t^h , τ_t^c , τ_t^l , τ_{t+1}^k could be set to an arbitrary value, and the optimal policy could still be implemented by choosing the remaining taxes appropriately. This is true not only in steady state, but also along a transition path. This might lead us to think that introducing an additional constraint on what the values of these taxes can take will not change the optimal solution, but this will not be the case.

If human capital investments cannot be observed separately from final consumption, then we cannot tax these two kinds of expenditure at different rates, which implies that $\tau_t^h = \tau_t^c$. Under this additional restriction, the household optimality conditions for human capital accumulation and labor become

$$\begin{aligned}\frac{u_{c,t-1}}{u_{c,t}} &= \beta n_t w_t \frac{1 - \tau_t^l}{1 + \tau_t^c} \\ -\frac{u_{n,t}}{u_{c,t}} &= h_t w_t \frac{1 - \tau_t^l}{1 + \tau_t^c}.\end{aligned}$$

In order to guarantee that the same ratio $(1 - \tau_t^l)/(1 + \tau_t^c)$ satisfies the two conditions, the planner's problem must now incorporate an additional restriction into its problem

$$\beta n_t u_{n,t} + h_t u_{c,t-1} = 0.$$

Incorporating this additional restriction associated to multiplier $\beta^t \delta_t$, the Lagrangean for the planner's problem becomes

$$L = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t) + \lambda c_t u_{c,t} - \delta_t [\beta n_t u_{n,t} + h_t u_{c,t-1}] \\ - \rho_t [c_t + g + k_{t+1} + h_{t+1} - F(k_t, l_t)] \end{array} \right\} - \lambda W_0$$

The first order conditions for this problem are

$$\begin{aligned}u_{c,t} + \mu(u_{c,t} + c_t u_{cc,t}) - \rho_t &= \delta_{t+1} h_{t+1} u_{cc,t} \\ u_{n,t} + \rho_t F_{l,t} h_t &= \delta_t (\beta u_{n,t} + \beta n_t u_{nn,t}) \\ \rho_{t+1} \beta F_{k,t+1} - \rho_t &= 0 \\ \rho_{t+1} \beta F_{l,t+1} n_t - \rho_t &= \delta_{t+1} u_{c,t}.\end{aligned}$$

The multiplier δ_t will remain positive, even in steady state. If this was not the case, then the labor wedge would be different from the human capital wedge. To see this, notice that with $\delta = 0$, the optimality condition for labor in steady state would be

$$-\frac{u_n}{u_c F_l h} = 1 + \mu(1 - \sigma) < 1,$$

whereas the wedge for human capital accumulation would be given by

$$\frac{1}{\beta F_l n} = 1,$$

which clearly violates the constraint associated with δ .

In steady state ρ must be constant and the choice of physical capital will remain undistorted. However, this will no longer be the case for human capital, since $\delta > 0$. With constant consumption taxes, optimal taxes will be given by

$$1 - \tau^k = 1 \text{ and } \frac{1 - \tau^l}{1 + \tau^c} = \frac{1 - \tau^l}{1 + \tau^h} < 1.$$

From the previous section, when differential taxes were allowed on final consumption and investment expenditures, optimal taxes were instead given by

$$1 - \tau^k = \frac{1 - \tau^l}{1 + \tau^h} = 1 \text{ and } \frac{1 - \tau^l}{1 + \tau^c} < 1.$$

When human capital expenditures were observable, it was optimal not to tax either type of capital in the long run and only tax labor. If they are not observable, the tax on human capital must be the same as the tax on labor, and will be positive, even in the long run.

4.4 Concluding Remarks

If the government cannot distinguish between final consumption and human capital expenditures, then it is optimal to tax both labor and human capital in the long run. Whether physical capital should be taxed or subsidized depends on its complemen-

parity with human capital versus raw labor. This is true even if consumption taxes are allowed.

Our results differ from those in Jones, Manuelli, and Rossi (1997) and Judd (1999), who find that human capital should not be taxed in the long run. The reason for this is that they assume that human capital expenditures are observable, which allows the government to subsidize them, and thus choose a different wedge for the intratemporal labor-consumption decision and the intertemporal human capital decision.

4.5 Appendix

So far we have assumed that the production function $F(k_t, l_t)$ has constant returns to capital and effective labor. Furthermore, since effective labor was given by $l_t = h_t n_t$, the production function had globally increasing returns to scale. Assume now that effective labor is actually a CRS function of human capital and raw labor given by $l_t = f(h_t, n_t)$, so that the production function $F(k_t, f(h_t, n_t))$ exhibits constant returns to scale in all production factors. Assume f is increasing and concave.

The new implementability condition is given by

$$\sum_{t=0}^{\infty} \beta^t (c_t u_{c,t} + n_t u_{n,t}) = W_0.$$

As before, if the government cannot distinguish between final consumption and human capital investment, it has to meet the following constraint

$$\beta f_{h,t} u_{n,t} + f_{n,t} u_{c,t-1} = 0.$$

The Lagrangean for the planner's problem is now given by

$$L = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t) + \lambda [c_t u_{c,t} + n_t u_{n,t}] - \delta_t [\beta f_{h,t} u_{n,t} + f_{n,t} u_{c,t-1}] \\ -\rho_t [c_t + g + k_{t+1} + h_{t+1} - F(k_t, f(h_t, n_t))] \end{array} \right\} - \lambda W_0.$$

The first order conditions for this problem are

$$\begin{aligned}
u_{c,t} + \mu(u_{c,t} + c_t u_{cc,t}) - \rho_t &= \delta_{t+1} f_{n,t+1} u_{cc,t} \\
u_{n,t} + \mu(u_{n,t} + n_t u_{nn,t}) + \rho_t F_{l,t} f_{n,t} &= \delta_t [\beta f_{h,t} u_{nn,t} + \beta f_{hn,t} u_{n,t} + f_{nn,t} u_{c,t-1}] \\
\rho_{t+1} \beta F_{k,t+1} - \rho_t &= \delta_{t+1} [\beta f_{hk,t+1} u_{n,t+1} + f_{nk,t+1} u_{c,t}] \\
\rho_{t+1} \beta F_{l,t+1} f_{h,t+1} - \rho_t &= \delta_{t+1} [\beta f_{hh,t+1} u_{n,t+1} + f_{nh,t+1} u_{c,t}].
\end{aligned}$$

We will now interpret these conditions with and without observable investment in human capital.

4.5.1 Observable Investment in Human Capital

When the government can observe investment in human capital the constraint associated with δ does not have to be met, which means that its multiplier is always zero. Given this, it is straightforward to see from the first order conditions for the optimal plan that neither type of capital should be distorted in the long run. On the other hand, the wedge between labor and consumption will remain distorted in the long run even if we do not impose that $\sigma > 1$ since now taxes will be given by

$$\frac{1 - \tau^l}{1 + \tau^c} = \frac{1 + \mu(1 - \sigma)}{1 + \mu(1 + \gamma)} < 1.$$

Thus, in this economy, if we choose consumption taxes to be zero, we can implement the optimal long run solution using zero capital taxes, positive labor taxes, and a subsidy on human capital expenditures that exactly offsets the distortion caused by the labor taxes on human capital accumulation.

Furthermore, notice that for this utility function, labor taxes are constant starting in period one, and capital taxes are zero starting in period two, so that there is no initial accumulation of assets that allows the government to sustain zero capital taxes in the long run. The same was true in the case of increasing returns to scale.

4.5.2 Unobservable Investment in Human Capital

As before, if human capital is unobservable, then it will be optimal to distort the human capital wedge in the long run, which must be the same as the intratemporal wedge between labor and consumption.

The optimal long run taxes on human capital (assuming zero consumption taxes) can be written as

$$\frac{\tau^l}{1 - \tau^l} = \beta F_l f_h - 1 = \frac{\delta}{\rho} u_c f_n \left[\frac{f_{nh}}{f_n} - \frac{f_{hh}}{f_h} \right].$$

The tax on human capital is always positive since $f_{nh} > 0$ and $f_{hh} < 0$.

Physical capital will not be taxed since the relative productivity of human capital and labor does not depend on the capital level

$$\frac{\tau^k}{1 - \tau^k} = \beta F_k - 1 = 0.$$

4.5.3 Taxation of Physical Capital

So far we have assumed that the production function can be written as $F(k_t, f(h_t, n_t))$, which assumes that the level of capital does not influence the relative productivity of human capital and labor. Assume now instead that the production function is given by $G(k_t, h_t, n_t)$, which has constant returns to scale with respect to the three factors of production. Firms can observe the contribution of each individual factor, so that equilibrium factor prices are given by

$$r_t^k = G_{k,t} \quad r_t^h = G_{h,t} \quad r_t^n = G_{n,t}.$$

With constant returns to scale in the production function, the households' budget constraint becomes

$$\sum_{t=0}^{\infty} p_t [c_t(1 + \tau_t^c) + k_{t+1} + (1 + \tau_t^h)h_{t+1} - (1 - \tau_t^k)r_t^k k_t - (1 - \tau_t^l)(r_t^n n_t + r_t^h h_t)] = p_0 b_0.$$

As in the previous section, the implementability condition is given by

$$\sum_{t=0}^{\infty} \beta^t (c_t u_{c,t} + n_t u_{n,t}) = W_0,$$

and the following constraint has to be met if the government cannot distinguish between final consumption and human capital investment

$$\beta G_{h,t} u_{n,t} + G_{n,t} u_{c,t-1} = 0.$$

The Lagrangean for the planner's problem is now given by

$$L = \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} u(c_t, n_t) + \lambda [c_t u_{c,t} + n_t u_{n,t}] - \delta_t [\beta G_{h,t} u_{n,t} + G_{n,t} u_{c,t-1}] \\ - \rho_t [c_t + g + k_{t+1} + h_{t+1} - G(k_t, h_t, n_t)] \end{array} \right\} - \lambda W_0.$$

Under this conditions, the first order condition for capital accumulation becomes

$$\rho_{t+1} \beta G_{k,t+1} - \rho_t = \delta_{t+1} [\beta G_{hk,t+1} u_{n,t+1} + G_{nk,t+1} u_{c,t}],$$

which means that when investment in human capital is not observable and δ is positive, the steady state level of capital taxes (assuming zero consumption taxes) is given by

$$\frac{\tau^k}{1 - \tau^k} = \beta G_k - 1 = \frac{\delta}{\rho} u_c G_n \left[\frac{G_{nk}}{G_n} - \frac{G_{hk}}{G_h} \right].$$

Thus, whether physical capital is taxed or subsidized depends on the relative magnitude of G_{nk}/G_n versus G_{hk}/G_h .

The reason for this is that since the constraint associated with δ is preventing labor taxes from being set optimally, it is optimal to also use capital taxation to reduce this distortion. This mechanism is similar to that in Correia (1996), where there were two kinds of labor, one of which could not be taxed (which means that its tax could not be set optimally), thus creating a reason for capital taxation, even in the long run.

For the case where the production function can be written as $G(k_t, h_t, n_t) =$

$F(k_t, f(h_t, n_t))$, the two effects exactly offset each other, so that it is optimal to have zero physical capital taxes in the long run.

Bibliography

- [1] ABREU, D. (1988) "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383-397.
- [2] ABREU, D., D. PEARCE, AND E. STACHETTI (1990) "Towards a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041-1064.
- [3] ACEMOGLU, D., M. GOLOSOV, AND A. TSYVINSKI (2006) "Markets versus Governments: Political Economy of Mechanisms," mimeo.
- [4] AGUIAR, M., M. AMADOR, AND G. GOPINATH (2006) "Efficient Expropriation: Sustainable Fiscal Policy in a Small Open Economy" mimeo.
- [5] AIYAGARI, R., A. MARCET, T. SARGENT, AND J. SEPPALA (2002) "Optimal Taxation without State-Contingent Debt," *Journal of Political Economy*, 110, 1220-1254.
- [6] ALBANESI, S. (2006) "Optimal Taxation of Entrepreneurial Capital with Private Information," *NBER* working paper.
- [7] AZZIMONTI-RENZO, M., P. SARTE, AND J. SOARES (2006) "Optimal Policy and (the lack of) Time Inconsistency: Insights from Simple Models," mimeo.
- [8] BANKS, J. P. DIAMOND, AND J. MIRRLEES (2007) "The Base for Direct Taxation," *Mirrlees Report*, draft.
- [9] BENHABIB, J. AND A. RUSTICHINI (1997) "Optimal Taxes without Commitment," *Journal of Economic Theory*, 77, 231-259.

- [10] BENVENISTE, L. AND J. SCHEINKMAN (1979) "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, 47, 727-732.
- [11] BOVENBERG, A AND B. JACOBS (2005) "Redistribution and Education are Siamese Twins" *Journal of Public Economics* 89.
- [12] BRETT AND WEYMARK (2003) "Financing Education using Optimal Redistributive Taxation" *Journal of Public Economics* 87.
- [13] BROADWAY, MARCEAU AND MARCHAND (1996) "Investment in Education and the Time Inconsistency of Redistributive Tax Policy" *Economica*, New Series, Vol. 63, No. 250.
- [14] CARD, D. (2001) "Estimating the Returns to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, Vol. 69, No. 5.
- [15] CHAMLEY, C. (1986) "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54, 607-622.
- [16] CHARI, V.V. AND P. KEHOE (1990) "Sustainable Plans," *The Journal of Political Economy*, 98, 783-802.
- [17] CHARI, V.V. AND P. KEHOE (1993a) "Sustainable Plans and Debt," *The Journal of Economic Theory*, 61, 230-261.
- [18] CHARI, V.V. AND P. KEHOE (1993b) "Sustainable Plans and Mutual Default," *The Journal of Economic Studies*, 60, 175-195.
- [19] CORREIA, I. (1996) "Should Capital Income be Taxed in the Steady State?," *Journal of Public Economics*, 60, 147-151.
- [20] COSTA, C. AND L. MAESTRI (2004) "The Risk Properties of Human Capital and the Design of Government Policies".
- [21] DIAMOND, P. (1989) "Optimal Income Taxation: an example with a U-shaped pattern of Optimal Marginal Taxes" *The American Economic Review*, Vol. 88, No. 1.

- [22] DIAMOND, P. AND J. MIRRLEES (2002) "Optimal Taxation and the Le Chatelier Principle" MIT working paper.
- [23] DOMINGUEZ, B. (2006) "Public Debt and Optimal Taxes without Commitment," forthcoming in *Journal of Economic Theory*.
- [24] FERNANDEZ-VILLAYERDE, J. AND A. TSYVINSKY (2002) "Optimal Fiscal Policy in a Business Cycle Model without Commitment," mimeo.
- [25] GOTTLIEB AND MOREIRA (2003) "Should educational policies be regressive?"
- [26] GROCHULSKI AND PISKORSKI (2005) "Optimal Wealth Taxes with Risky Human Capital" Federal Reserve of Richmond WP, No. 05-13.
- [27] HARE AND ULPH (1979) "On Education and Distribution" *The Journal of Political Economy*, Vol. 87, No. 5, Part 2.
- [28] JONES, L., MANUELLI, R., AND P. ROSSI (1993) "Optimal Taxation in models of Endogenous Growth," *Journal of Political Economy*, 101, 485-517.
- [29] JONES, L., MANUELLI, R., AND P. ROSSI (1997) "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73, 93-117.
- [30] JUDD, K. (1985) "Redistributive Taxation in a Simple Foresight Model," *Journal of Public Economics*, 28, 59-83.
- [31] KAPICKA, M. (2006) "Optimal Income Taxation with Human Capital Accumulation and Limited Recordkeeping," *Review of Economic Dynamics*, 9, 612-639.
- [32] KYDLAND, F. AND E. PRESCOTT (1977) "Rules rather the Discretion: The inconsistency of Optimal Plans" *Journal of Political Economy*, 85, 473-491.
- [33] KLEIN, P. AND J. RIOS-RULL (2003) "Time-Consistent Optimal Fiscal Policy," *International Economic Review*, 44, 1217-1246.
- [34] KLEIN, P., P. KRUSELL, AND J. RIOS-RULL (2004) "Time-Consistent Public Expenditures," CEPR Discussion Paper 4582.

- [35] KLEIN, P., V. QUADRINI, AND J. RIOS-RULL (2005) "Optimal time-consistent taxation with international mobility of capital," mimeo.
- [36] KLEIN, P., P. KRUSELL, AND J. RIOS-RULL (2006) "Time-Consistent Public Policy," mimeo.
- [37] KOCHERLAKOTA, N. (1996) "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63, 595-609.
- [38] LJUNGQVIST, L. AND T. SARGENT (2004) *Recursive Macroeconomic Theory*, 2nd edition.
- [39] LUCAS, R., JR. AND N. STOKEY (1983) "Optimal Fiscal and Monetary Policy in an Economy without Capital" *Journal of Monetary Economics*, 12, 55-93.
- [40] LUCAS, R., JR. AND N. STOKEY WITH E. PRESCOTT (1989) *Recursive Methods in Economic Dynamics*.
- [41] MARCET, A. AND R. MARIMON (1992) "Communication, Commitment, and Growth," *Journal of Economic Theory*, 58, 219-249.
- [42] MIRRLEES, J. (1971) "An Exploration in the Theory of Optimum Income Taxation" *The Review of Economic Studies*, Vol. 38, No. 2.
- [43] MIRRLEES, J. (1986) "The Theory of Optimal Taxation" *Handbook of Mathematical Economics*, edited by K.J. Arrow and M.D. Intrilligator, Vol. III, ch. 24.
- [44] PLELAN, C. AND E. STACCHETTI (2001) "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 69, 1491-1518.
- [45] RAY, D. (2002) "Time Structure of Self-Enforcing Agreements," *Econometrica*, 70, 547-582.
- [46] WERNING, I. (2003) "Standard Dynamic Programming for Aiyagari, Marcet, Sargent and Seppala," Lecture Note.