Essays on the Political Economy of Taxation in Dynamic Settings

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Juan Facundo Piguillem

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Larry Jones

August, 2009

Acknowledgments

I'd like to thank my advisor Larry Jones for his continuous help and support. I also thank Anderson Scheneider, V.V. Chari, Katya Kartashova, Hakki Yazici, Narayana Kocherlakota, Pricila Maziero, Ellen R. McGrattan, Chris Phelan, Songying Fang, Warren Weber, participants at the Midwest Macro conference in Minneapolis, the SED meetings in Boston and participants at the Federal Reserve Bank of Minneapolis bag lunch for comments and suggestions.

Dedication

This dissertation is dedicated to my friends.

Contents

1 List of Tables

List of Tables

2 List of Figures

List of Figures

Part I

Introduction

Among the possible reasons for how people vote in democratic regimes, the magnitude of taxes and their potential redistribution effect would seem to be a first order issue for most voters. We think that a better understanding of the conflict of preferences over collective objects like taxes represents an important subject in the public economics area. Moreover, it is reasonable to think that the tax's properties would be potentially affected by the assumptions on the politico-economic process. Thus, even though the spirit of every democratic institutional design is to approximate the true preferences of society in the best possible way, different institutional designs would in general imply different levels and stochastic properties for different kind of taxes (which do not necessarily correspond to the society's preferences). In addition, properly including all the relevant details involved in any particular institutional design, if possible, may render a model quickly intractable. In spite of this, most of the work that analyses inefficiencies due to political constraints has followed the route of making ex ante assumptions about the forms of institutions.

The last approach typically requires making very specific assumptions about the institutions that are used to generate policies (e.g., specific game theoretic models of voting over a restricted set of tax instruments) along with a variety of other imperfections. This makes it hard to interpret the results since it is not clear if the properties of the policies singled out as equilibria are chosen due to the specific institutional arrangements assumed or due to the imperfections added. This leaves too much room for the researcher to choose, arbitrarily, which are the key components of the institutions to be considered. In addition, minor variations in the components chosen or its main characteristics could have an important impact on the conclusions.

An alternative path, more in the tradition of classic Public Choice literature, is to appeal to The Median Voter Theorem. This is a useful device to aggregate preferences that delivers a simple and strong result: the equilibrium policy is the most preferred by the median voter (Condorcet winner). Unfortunately, it is well known that the Median Voter Theorem only holds under very specific assumptions that are not satisfied in general, as discussed in Kramer (1973). Because of these restrictions, only a few results are available in the public economics literature. In general, such results arise from assumptions that imply indirect preferences over policies being single peaked. However, when the result holds its policy implications are robust. [1] Mainly motivated by these issues we consider that following the last path could prove to be a fruitful and important approach to analyzed Political Economy problems in Macroeconomics.

In this thesis we focus on the voting decisions of heterogeneous agents about redistributive fiscal policy that is also used to finance a stochastic stream of gov-

[1] In the sense that it is easy to find a variety of institutional designs that would implement the prescribed policy

ernment expenditures. In Chapter 2, they vote at time zero (once and for all) on a sequence of lump-sum transfers (or payments) and also on marginal taxes on both capital and labor income, while in Chapter 3 they vote sequentially. Disagreement of individual preferences over capital income taxes will be motivated by differences in initial wealth levels. The conflict over labor taxes is introduced through a kind of heterogeneity that we think is particularly important in the real world: people are born with different abilities to turn effort into input. In all other aspects, the environment analyzed borrows their main components from the Stochastic Neoclassical Growth Model with complete markets. Thus, we build on the work of Bassetto and Benhabib (2006) (henceforth B&B) adding leisure decision and labor skills heterogeneity.

Since people are heterogeneous with respect to different economic variables, different agents will eventually disagree about the size of specific taxes and the extent of the redistribution. Furthermore, agents decide their vote based, among other things, on a "package" of taxes. That means people with different characteristics will disagree not only about their best choice of a specific tax, but also about the package including all taxes. In this type of economy in which redistribution is allowed, it is reasonable to think that poor households endowed with low labor skills will prefer a package including high taxes on both labor and capital income. On another hand, rich individuals endowed with high skills will prefer a package having low taxes. But is there a specific fiscal policy that can not be defeated by any other policy through majority voting? If such policy does exist, what are capital and labor income taxes? In addition, how does individual heterogeneity shape such a policy?

Chapter 2 answers the above questions when agents are constrained to vote only once, at the beginning of period zero, and never again. Specifically, in Proposition 1, assuming balanced growth preferences, we give two different sets of sufficient conditions for the existence of a Condorcet winner. The first part of Proposition 1 assumes that there is no heterogeneity in labor skills, although agents value leisure. In this case, we show that for every distribution of initial endowments, the best fiscal policy for the agent with the median endowment is preferred to any other policy by at least half of the individuals in the economy. The second part of the proposition considers heterogeneity in both labor skills and initial wealth. Then, if the distribution of initial capital endowments is an affine function of the distribution of labor skills, again the most preferred time path of policies is preferred by a majority.

Regarding the characterization of the Condorcet Winner, and under the assumptions of Proposition 1, it follows that given any fiscal policy the indirect utility for the median agent can be decomposed into two parts: a redistributive component and an efficiency component. The redistributive component depends directly on the skewness of the distribution of skills and it is increasing in the distortions yielded by both capital and labor taxes. The efficiency component is given by the value of the mean agent's indirect utility and it is decreasing in any positive distortion. The most preferred tax schedule for the median type balances these two components. As in B&B, we show that capital income taxes will be either zero or at the upper bound in any period and state, with at most one period in between. In addition, Proposition 3 shows that marginal taxes on labor income depend directly on the absolute value of the distance between the median and the mean value of the labor productivity distribution in the economy.

Finally, we test the theory of Chapter 2 empirically. A calibrated version of the model without initial wealth inequality is used to account for the observed increasing trend in both labor income inequality and average tax on labor in the U.S. in the last decades. The model does a good job of fitting both the increasing trend and the levels of labor, and also of matching some short run co-movements. The model accounts for twice the growth in labor taxes observed in the period 1962-2001. It also yields correlations between labor taxes and other aggregate variables with the right sign and similar magnitudes.

The results of Chapter 2 were derived under the strong assumption that agents vote only once. Are those results still true when agents can vote in every period? (i.e. when they can revise their past decisions). This is important because the equilibrium capital income tax of Chapter 2 is time inconsistent. That is, it prescribes the maximal possible tax rate in the initial period and then converges to zero over time. Thus, if agents have the possibility to choose again, when zero capital taxation should be in effect, they would revise their decision and again choose the highest possible capital income tax. Chapter 3 deals with this problem. Specifically, it analyzes the role of reputation in influencing political outcomes about fiscal policy. In each period society decides the current fiscal policy according to a variation of the majority voting concept in Bernheim and Slavov (2009). The basic feature of this equilibrium concept is that current choices of both taxes and individual decisions imply different continuation policies. The first result is the existence of a Dynamic Condorcet Winner which coincides with the preferences of the median voter, the agent holding the median wealth at time zero. The second result is a characterization of equilibrium outcomes when preferences are logarithmic, the production function is Cobb-Douglas and there is full depreciation of capital. The most preferred competitive equilibrium for the median type with commitment can be sustained as an equilibrium outcome for high enough discounting.

Part II

Heterogeneous Labor Skills, The Median Voter and Labor Taxes[2]

3 Introduction

Even though a great extent of work has been done to analyze the characteristics of taxes determined by politico-economic process, little is known about the properties of labor taxes. One way to construct a positive theory about fiscal policy is to assume that agents optimally choose the policies that will take place in the future. This optimal choice may, in part, be motivated by a first order issue in elections: the redistributive effects of the different policies. Following this idea, we study a class of dynamic model economies with heterogeneous agents where the only political institution is the pursuit of consensus. Agents vote, once and for all, at the beginning of time on sequences of redistributive taxes on capital and labor. Building on the work of Bassetto and Benhabib (2006) (henceforth B&B), we derive a median voter theorem for this class of economies. We use this theorem to describe the properties of the equilibrium tax sequences. The theorem gives one, precise, statement of the form that redistribution considerations take in determining policy.

As in B&B, we add fiscal policy that allows for redistribution in the standard

[2]This chapter is coauthored with **Anderson Schneider**

neoclassical growth model. We go beyond their paper by adding leisure choice and stochastically evolving labor productivities. Along with this, we also add to their framework marginal taxes on labor income. Individual's disagreement over capital income taxes is motivated by differences in initial wealth levels. The conflict about labor taxes is given by heterogeneous labor skills: agents have different abilities to turn effort into effective labor, which is rented to firms in the market. Although we do not model any voting process explicitly, consider the following situation: at time zero, before the economy starts, all possible sequences representing different fiscal policies are analyzed, and a consensus should be reached through sincere majority voting. By sincere we mean an agent would vote for fiscal policy A against policy B if she prefers A to B. A natural question then arises: is it actually possible to reach such a consensus? Or in other words, is there any policy that precludes the existence of Condorcet cycles when sincere voting is in place? If such a policy does exist, what are the capital and labor income taxes implied and how does individual heterogeneity shape this policy?

In proposition 1, assuming balanced growth preferences, we give two different sets of sufficient conditions for the existence of a Condorcet winner in this type of environment. The first part of proposition 1 assumes that there is no heterogeneity in labor skills, although agents value leisure. In this case, we show that, independent of the distribution of initial endowments, the best fiscal policy (consisting of a full time path of both labor and capital income tax rates) for the agent with the median endowment is preferred to any other policy by at least half of the individuals in the economy. The second part of the proposition considers heterogeneity in both labor skills and initial wealth. If the initial capital endowment is an affine function of labor skills, then again, the most preferred time path of policies is preferred by a majority.

The proof of the consensus result relies on a characterization of indirect preferences over fiscal policies that is of independent interest. Although proposition 1 can be thought in terms of fiscal policies, it is actually stated in terms of implementable allocations: those that can be decentralized as a competitive equilibrium. Under complete markets, if agents have the same balanced growth utility function, individual allocations can be expressed as a constant share of their aggregate counterparts. These shares are functions of both types and aggregate allocations. Moreover, the indirect preferences, as a function of types, inherit the properties of these share functions. Then, we show that for any two fiscal policies for which a competitive equilibrium exists, the indirect preferences can cross at most once in the space of types, delivering the result.

Our second contribution concerns the characterization of the Condorcet Winner. It follows that the indirect utility for the median type over any fiscal policy can be decomposed into two parts: a redistributive component and an efficiency one. The redistributive component depends directly on the skewness of the distribution of skills and it is increasing in the distortions yielded by both capital and labor taxes (together with bigger transfers). The efficiency component is given by the value of the mean type's indirect utility and it is decreasing in any positive distortion. The most preferred tax schedule for the median type balances these two components. As in B&B, we show that capital income taxes will be either zero or at the upper bound in any period and state, with at most one period in between.

In addition, Proposition 3 shows that marginal taxes on labor income depend directly on the absolute value of the distance between the median and the mean value of the productivity distribution in the economy.

The results are extended to the case that skills evolve stochastically over time keeping constant the ranking among agents. Again labor taxes depend directly on the skewness of the distribution. However, the final outcome is ambiguous - this relationship could be either increasing or decreasing. We study a numerical example and show that the model without capital accumulation can generate either procyclical or counter-cyclical taxes, depending on how the dispersion of the distribution of individual productivities changes along the business cycle. This may offer an answer to the question posed by Alesina, Campante and Tabellini (2008), related to the empirical observation that fiscal policy is often procyclical in developing countries and counter-cyclical in developed ones.

Finally, a calibrated version of the model without initial wealth inequality is used to check if the theory can account for the observed increasing trend in both labor income inequality and average tax on labor in US in the last decades. The calibration for the skill process is done using data on wages from Eckstein and Nagypal (2004). The model does a good job on fitting both the increasing trend and the levels of labor taxes in the last decades, and also on matching some short run co-movements. The model accounts for twice as much of the growth in labor taxes observed in the period 1962-2001. It also yields a negative correlation between taxes and aggregate labor, in line with the data.

We view the results regarding labor taxes as a neat characterization of an important component of fiscal policy. Most of the work that analyzes inefficiencies due to political constraints, has followed the route of making strong, ex ante, assumptions about the forms of institutions. The difficulty with this approach is that it typically requires making very specific assumptions about the institutions that are used to generate policies (e.g., specific game theoretic models of voting over a restricted set of tax instruments) along with a variety of other imperfections. This makes it hard to interpret the results since it is not clear if the properties of the policies singled out as equilibria are chosen due to the specific institutional arrangements assumed or due to the imperfections added.

As we mentioned before, our model gives an extension to the median voter result presented in Bassetto and Benhabib (2006).[3] Other than the already highlighted differences in the characteristics of the physical environment, our results depend

[3] Important contributions on median voter results and its connection with fiscal policy include Meltzer and Richard (1981), Alesina and Rodrik (1994), Persson and Tabellini (1994), among others.

on the assumption of balanced growth preferences defined over consumption and leisure. On the other hand, although B&B do not consider leisure choice, a more general class of Gorman aggregable preferences is analyzed.

Besides the work by Bassetto and Benhabib (2006), five other papers deserve special mention. Werning (2006) considers the same physical environment as here and analyzes the Ramsey outcome when the government uses fiscal policy for redistribution and to finance an exogenous stream of expenditures. The possibility of non distortionary taxation is not ruled out ex-ante, nevertheless distortions emerge in the decentralized solution, regardless of the welfare weights used by the government. We find similar results, although we obtain a more specific characterization of labor taxes. This feature comes partially from the fact that the median voter solution uses welfare weight equal to one for the median type. In the case that agents have stochastic labor skills, a numerical exercise in his paper shows that the implied labor taxes from a Utilitarian Ramsey problem comove with the distribution of skills. The author does not provide a numerical solution calibrated to the US economy.

Azzimonti, De Francisco, and Krusell (2008) also analyze majority voting over marginal taxes on labor income. Since their environment does no consider both aggregate uncertainty and capital accumulation, the best sequence of labor taxes for each type in the economy can be characterized by two numbers (taxes in the first two periods). A median voter result is provided in the case where either there is heterogeneity in the initial wealth only or in the labor skills.

Krusell and Rios-Rull (1999) consider an environment similar to ours but voting takes place periodically, taxes on capital and labor income are constrained to be equal and only future taxes can be changed. A Markov stationary equilibrium is solved numerically. The stationary equilibrium exhibits positive distortions. As in this paper, the level of income taxation depends on the skewness of income distribution. Since their paper consider a marginal tax on income, results about labor taxes are not provided.[4]

Regarding the empirical results, Chari, Christiano, and Kehoe (1994) analyze the quantitative implications of optimal fiscal policy in a dynamic model with homogenous agents. We emphasize that they consider the same class of balanced growth preferences. Using different calibrated versions of the model, they found that labor taxes are essentially constant over the business cycle, although labor taxes inherit the stochastic properties of the exogenous shocks (productivity and government spending). Finally, Corbae, D'Erasmo, and Kuruscu (2009) use a recursive political economy model, as in Krusell and Rios-Rull (1999), to evaluate how much the increase in wage inequality in the period 1979-1996 can account for the relative increase in both transfers to low earnings quintiles and effective tax rates for higher quintiles. They assume idiosyncratic labor skills shocks and incomplete markets. The paper uses a median voter result by checking numerically that prefer-

[4]Azzimonti, De Francisco, and Krusell (2008) provide an analytical characterization of time-consistent Markov-perfect equilibria in an environment similar to Krusell and Rios-Rull (1999), but individual heterogeneity is restricted to initial wealth.

ences over one period income taxes (on and off-path) are single-peaked. They found that the model predicts about half of the increase in redistribution to lowest wage quintiles, and also it overpredicts the average effective tax rate.

The Chapter proceeds as follows. Section 2 describes the environment. Section 3 characterizes the competitive equilibrium given a fiscal policy. In section 4 we construct the proof for the consensus result. Section 5 characterizes the Condorcet winner, while section 6 considers stochastic skills. Section 7 solves numerically the model without capital accumulation, and the last section concludes.

4 The Economy with Constant Skills

There is a continuum of agents indexed by the labor skill parameter $\theta \in \Theta \equiv [\underline{\theta}, \overline{\theta}]$ with $\underline{\theta} > 0$. Later we relax this assumption, allowing for stochastic labor skills. The distribution of θ is represented by the p.d.f. $f(\cdot)$ and the median type is denoted by $\theta^m \leq \int_\Theta \theta f(\theta) d\theta = 1$ by assumption.[5]

Uncertainty is driven by the public observable state $s_t \in S$, where S is finite. It potentially affects the efficient production frontier. Let $s^t = (s_0, ..., s_t)$ be the history of shocks up to time t and $\Pr(s^t)$ its marginal probability. Assume that $\Pr(s_0 = \overline{s}) = 1$ for some $\overline{s} \in S$.

The output at time t is produced by competitive firms using capital and efficient labor. The resources constraint for each pair (t, s^t) is

[5]The median voter result presented later does not depend on this skewness assumption.

$$C(s^t) + K(s^t) \leq F\Big(L(s^t), K(s^{t-1}), s^t\Big) + (1-\delta)K(s^{t-1}) \tag{1}$$

where the function $F(\cdot)$ is assumed to be homogeneous of degree one in both capital and labor for all s^t.

Remark: We could consider an exogenous stream of government expenditures without changing the main results. But since our concern is mainly related to redistribution, the restriction of zero government consumption avoids dealing with valuations of the benefits of positive marginal taxes net of the distortions in financing government expenditures. In the numerical exercise presented later we consider exogenous government spending.

Each agent has an endowment of one unit of time in each period and state. Using l/θ units of its time agent type θ produces l units of efficient labor that is rented to the firms. If agent type θ consumes the stream $\{c_t, 1 - l_t/\theta\}_{t=0}^{\infty}$ of consumption and leisure, then its total discounted utility is given by:

$$\sum_{t=0}^{\infty} \beta^t u\Big(c(s^t,\theta), 1 - l(s^t,\theta)/\theta\Big)$$

where we assume balanced growth preferences, i.e., $u(c, 1-l) \equiv \frac{\left[c^{\alpha}(1-l)^{1-\alpha}\right]^{1-\sigma}}{1-\sigma}$ if $\sigma \neq 1$, or $\alpha \log(c) + (1-\alpha)\log(1-l)$ if $\sigma = 1$.

In the initial period, agent type θ is endowed with $W_{-1}(\theta) \equiv k_{-1}(\theta) > 0$ units of wealth consisting of capital stock. Later we shall impose conditions on the initial wealth distribution.

In each period the government levies an affine tax schedule on labor income given by $\tau_l(s^t)w(s^t)l(s^t;\theta)+T(s^t)$, where $w(s^t)$ are the wage payments and the lump-sum tax $T(s^t)$ is potentially used for redistribution. Notice that the tax schedule is not individual specific.

The government taxes capital returns net of depreciation at rate $\tau_k(s^t) \in [0, \overline{\tau}]$. For the type of wealth distribution that we analyze later, the lower bound will never bind. The upper bound on capital taxes is a technical condition required in order to guarantee that the best allocation for the median type exists. In order to reduce the arbitrariness of such an exogenous upper bound, we choose $\overline{\tau} = 100\%$. In this way the maximum levy corresponds to a loss of the full return net of depreciation.

Profit maximization by the firms determines the rental prices. Given a tax sequence, prices, and initial endowments, under complete markets agent type θ chooses his individual allocation in order to maximize utility subject to the budget constraint:

$$\sum_{t,s^t} p(s^t)\Big(c(s^t;\theta)+k(s^t;\theta)\Big) \leq \sum_{t,s^t} p(s^t)\Big((1-\tau_l(s^t))w_t(s^t)l(s^t;\theta)+R(s^t)k(s^{t-1};\theta)\Big)-T \tag{2}$$

where $T \equiv \sum_{t,s^t} p(s^t)T(s^t)$ is the present value of the lump-sum taxes and $R(s^t) \equiv 1+(1-\tau_k(s^t))(r(s^t)-\delta)$.

Under the complete markets assumption the government budget constraint can be written as:

$$-T \leq \sum_{t,s^t} p(s^t)\bigg(\tau_l(s^t)w(s^t)L(s^t) + \tau_k(s^t)(r(s^t - \delta))K(s^{t-1})\bigg) \tag{3}$$

The usual definition for a competitive equilibrium follows:

Definition 1. *A competitive equilibrium given a tax schedule* $\{\tau_l(s^t), \tau_k(s^t), T(s^t)\}_t$ *is a sequence of prices* $\{w(s^t), p(s^t), r(s^t)\}_t$*, individual allocations* $\{c(s^t;\theta), l(s^t;\theta), k(s^t;\theta)\}_t$ *and implied aggregate allocations* $\{C(s^t), L(s^t), K(s^t)\}_t$ *such that:*

1. *Given after-tax prices,* $\{c(s^t;\theta), l(s^t;\theta), k(s^t;\theta)\}$ *maximizes utility subject to (2);*

2. $C(s^t) = \int_\Theta c(s^t;\theta)f(\theta)d\theta$, $L(s^t) = \int_\Theta l(s^t;\theta)f(\theta)d\theta$ *and* $K(s^t) = \int_\Theta k(s^t;\theta)f(\theta)d\theta$*;*

3. *Factor prices are equal to the marginal products;*

4. *The government budget constraint (3) holds; and*

5. *The resource constraint in (16) holds for every* s^t

5 Equilibrium Characterization

Here we characterize the economy given a fiscal policy for the log utility case. We use a characterization strategy similar to Werning (2006) and it will be very useful to prove the consensus result

[6]

The first order conditions with respect to individual consumption and output yields:

$$\begin{aligned}
\frac{\alpha\beta^t \Pr(s^t|s_0)}{c(s^t;\theta)} &= p(s^t)\lambda(\theta) && (4)\\
\frac{(1-\alpha)\beta^t \Pr(s^t|s_0)}{\theta - l(s^t;\theta)} &= p(s^t)(1-\tau_l(s^t))w_t(s^t)\lambda(\theta) && (5)
\end{aligned}$$

Let $\varphi(\theta) \equiv 1/\lambda(\theta)$, and $E(\varphi) \equiv \int_\Theta \varphi(\theta) f(\theta) d\theta$. Then integration over types in the expressions above yields the following equations determining after-tax prices:

$$\begin{aligned}
p(s^t) &= \frac{E(\varphi)\alpha\beta^t \Pr(s^t|s_0)}{C(s^t)} && (6)\\
p(s^t)(1-\tau_l(s^t))w_t(s^t) &= \frac{E(\varphi)(1-\alpha)\beta^t \Pr(s^t|s_0)}{1-L(s^t)} && (7)
\end{aligned}$$

We normalize $p_0 = \alpha/C_0$. This makes $E(\varphi) = 1$. Then individual allocations can be written as:

$$\begin{aligned}
c(s^t;\theta) &= \varphi(\theta)C(s^t) && (8)\\
1 - l(s^t;\theta)/\theta &= \varphi(\theta)\theta^{-1}[1-L(s^t)] && (9)
\end{aligned}$$

The other conditions for optimization in the problem faced by individual θ are

$$p(s^t) = \sum_{s^{t+1}} R(s^{t+1})p(s^{t+1}), \quad \text{and} \quad \lim_{t\to\infty} \sum_{s^t} p(s^t)k(s^t;\theta) = 0$$

[6]It turns out that the characterization in the logarithmic case is much simpler than in the general case (balanced growth preferences). All the proofs in the general case are shown in the Appendix A

Replacing the two conditions above, CE prices, and individual allocations expressed as a share of aggregates in the budget constraint for agent θ yields:

$$\varphi(\theta) = (1-\beta)\left[\widetilde{W}_0(\theta, T, \tau_{k0}) + \theta \sum_{t,s^t} \beta^t \Pr(s^t)\left[\frac{(1-\alpha)}{(1-L(s^t))}\right]\right] \tag{10}$$

where $\widetilde{W}_0(\theta, T, \tau_0) \equiv (\alpha/C_0)R_0 k_{-1}(\theta) - T$.

Since for each type the expression for $\varphi(\cdot)$ depends on the aggregate allocations and the tax schedule, the function can be rewritten as $\varphi(Z;\theta) \in \mathbb{R}_+$ where $Z \equiv (\{C(s^t), L(s^t), K(s^t)\}_{t\geq 0}, T, \tau_{k0})$. Let Z^∞ be the set of such sequences.

One can see from the construction of (10) that $\varphi(Z;\theta)/(1-\beta)$ represents the total value of expenditures on consumption and leisure expressed in efficient units for type θ. In equilibrium, this value should be equal to the value of the initial endowment plus the maximal present discounted value of his labor income, as equation (10) shows. If for some $\theta \in \Theta$ the value of $\varphi(Z;\theta)$ is equal to zero, then this individual can afford only zero consumption and zero leisure in all periods and states.

Remark: For future use, it is straightforward to replicate the computations above for the case in which the heterogeneity is only restricted to the initial endowments. In this case we would have individuals indexed by the initial capital endowment distributed according a pdf $f(\cdot)$ on $[\underline{k}, \overline{k}]$. The labor skills are given by the constant function $\theta(k_{-1}) = 1 \; \forall \; k_{-1} \in [\underline{k}, \overline{k}]$. In this case one can find that $\varphi(Z; k_{-1}) = (1-\beta)[\widetilde{W}_0(k_{-1}, T, \tau_0) + \sum_{s^t} \beta^t \Pr(s^t)(1-\alpha)/(1-L(s^t))]$.

Then, as in Werning (2006), we have the following:

Lemma 1. *$Z \equiv (\{C(s^t), L(s^t), K(s^t)\}_{t\geq 0}, T, \tau_{k0})$ is the aggregate allocation sequence (together with T and $\tau_{k0} \leq \overline{\tau}$) in an interior CE if and only if:*

1. *Z satisfy the resources constraint in (16) for all s^t;*

2. $\frac{1}{C(s^t)} \geq \beta E\left[\frac{1+(1-\overline{\tau})(F_k(s^{t+1})-\delta)}{C(s^{t+1})}|s^t\right]$ *for all s^t;*

3. *Evaluated at the aggregate allocations, the function $\varphi : Z^{\infty} \times \Theta \to \mathbb{R}_+$ given by (10) is such that $\varphi(Z;\theta) \in (0, \frac{1}{1-L(s^t)})$ for all $s^t, \theta \in \Theta$, and $L(s^t) \subset Z$. Furthermore, $E(\varphi) = 1$.*

Proof: Necessity comes from the reasoning above. Sufficiency is shown in the appendix.

The second condition comes from the upper bound on the capital tax rates. The last condition comes from the nonnegativity of consumption and the fact that leisure is bounded by the unit. It replaces the usual implementability conditions found in the Ramsey literature.

Depending on the restrictions on the distribution of the initial endowment, we could relax the third condition in the Lemma above to $\varphi(Z;\theta) \geq 0$. For example, this would be the case if initial endowments are non-decreasing in the skill level, making $\varphi(\cdot;\cdot)$ strict increasing in θ.

Because preferences are homothetic, Lemma 1 imples that, given taxes, two economies having different distributions of productivity types with the same mean and the same initial aggregate capital stock will have the same aggregate outcomes in equilibrium. Clearly, the distribution of φ in the economy will indeed depend on the distribution of skills and the assumptions on the initial endowments. Also notice that the distortions generated by marginal taxes are enclosed in the aggregates that determine the function φ.

As Chari and Kehoe (1999) have pointed out, the non-arbitrage condition $p(s^t) = \sum_{s_{t+1}} p(s^{t+1})R(s^{t+1})$ does not uniquely pin down the stochastic process for the capital tax rate.

Next we give a more intuitive representation for the function φ. Using (10) and the fact that $E(\varphi) = 1$, the constant of proportionality can be rewritten as:

$$\varphi(Z;\theta) = 1+(1-\beta)\left[\left(\widetilde{W}_0(\theta, T, \tau_{k0}) - \mathrm{E}(\widetilde{W}_0(\theta, T, \tau_{k0}))\right) + (\theta - \mathrm{E}(\theta)) \cdot UL(Z)\right] \geq 0$$

where $UL(Z) \equiv \sum_{t,s^t} \beta^t \Pr(s^t)\left[\frac{(1-\alpha)}{(1-L(s^t))}\right]$.

Therefore individuals that are wealthier than the average will have both individual consumption and leisure (measured in inefficient units) higher than the respective aggregates.

6 The Consensus Result

Let Ξ be the set of elements $Z \equiv (\{C(s^t), L(s^t), K(s^t)\}_{s^t}, T, \tau_{k0})$ that satisfy the conditions in Lemma 2.

We begin by analyzing preference orderings over elements of Ξ. As we have pointed out, all distortions generated by marginal taxes are already enclosed in the aggregate allocations. Given any $Z \in \Xi$, we can express the present discounted utility for each individual when aggregate allocations are given by Z. For agent type θ, denote this value by $V(Z;\theta)$. Then we have:

$$V(Z;\theta) = \frac{1}{1-\beta}\left[\log\left(\varphi(Z;\theta)\right) - (1-\alpha)\log(\theta)\right] + \\ \sum_t \beta^t \Pr(s^t)\left[\alpha\log(C(s^t)) + (1-\alpha)\log(1-L(s^t))\right] \quad (11)$$

The share of each individual can be rewritten as $\varphi(Z;\theta) = Bk_{-1}(\theta) + C\theta + D$, where B, C, and D are values that depend on $Z \in \Xi$. For a given $Z \in \Xi$ and associated $\varphi(\cdot)$, with some abuse of notation, let $J_\varphi(\theta)$ and $J(Z)$ be respectively the first and second term of (11).

Agent θ weakly prefers the allocation Z to $\hat{Z}$ if and only if

$$V(Z;\theta) \geq V(\hat{Z};\theta) \Longleftrightarrow J_\varphi(\theta) + J(Z) \geq J_{\hat{\varphi}}(\theta) + J(\hat{Z})$$

Let $S_{Z,\hat{Z}} = \left\{\theta : J_\varphi(\theta) - J_{\hat{\varphi}}(\theta) \geq J(\hat{Z}) - J(Z)\right\}$, that is, the set of agents that prefers Z to $\hat{Z}$.

As mentioned before, in this section we state the results for the logarithmic case. The proof for the more general Cobb-Douglas class of utility functions is shown in Appendix A. The general strategy of the proof presented below is similar to the one used to prove proposition 2 in Benhabib and Przeworski (2006).

Proposition 1. *: Assume balanced growth preferences and consider any $Z, \hat{Z} \in \Xi$.*

1. *Let heterogeneity be restricted **only** to the initial endowments, distributed according pdf $f(\cdot)$ on $[\underline{k}, \overline{k}]$ with mean K_{-1}. Denote k_{-1}^m be the the agent with the median wealth. If $k_{-1}^m \in S_{Z,\hat{Z}}$, then either $[\underline{k}, k_{-1}^m] \subseteq S_{Z,\hat{Z}}$ or $[k_{-1}^m, \overline{k}] \subseteq S_{Z,\hat{Z}}$.*

2. *Let agents be heterogeneous with respect to both labor skills and initial endowment. Also suppose that the initial endowment is an affine function of the skills. If $\theta^m \in S_{Z,\hat{Z}}$ then either $[\underline{\theta}, \theta^m] \subseteq S_{Z,\hat{Z}}$ or $[\theta^m, \overline{\theta}] \subseteq S_{Z,\hat{Z}}$.*

Proof. Due to its simplicity, we present here the proof for the logarithm case. In Appendix A we show the proof for the general case.

Take any $Z, \hat{Z} \in \Xi$; we shall show that $J_\varphi(\theta^m) - J_{\hat{\varphi}}(\theta^m) \geq J(\hat{Z}) - J(Z)$ implies $J_\varphi(\theta) - J_{\hat{\varphi}}(\theta) \geq J(\hat{Z}) - J(Z)$ for at least 50% of the agents. A sufficient condition for this to happen is the function $J_\varphi(\theta) - J_{\hat{\varphi}}(\theta)$ being monotone.

We start with the first statement. Given the remark after (10), heterogeneity only with respect to initial endowments allow us to write $\varphi(Z; k_{-1}) = Bk_{-1} + C + D$. Moreover, condition 3) in Lemma 1 implies $\varphi(Z; k_{-1}) = 1 + B(k_{-1} - K_{-1})$.

Given Z and $\hat{Z}$, without loss of generality assume $B > \hat{B}$. If $k_{-1}^m \geq K_{-1}$, then $J_\varphi(k_{-1}^m) \geq J_{\hat{\varphi}}(k_{-1}^m)$ and $J_\varphi(k_{-1}) \geq J_{\hat{\varphi}}(k_{-1})$ for all $k_{-1} \geq k_{-1}^m$. If $k_{-1}^m \leq K_{-1}$, then $J_\varphi(k_{-1}^m) \leq J_{\hat{\varphi}}(k_{-1}^m)$ and $J_\varphi(k_{-1}) \leq J_{\hat{\varphi}}(k_{-1})$ for all $k_{-1} \leq k_{-1}^m$. This proves the first statement.

Next, consider the case with heterogeneous labor skills: $\varphi(Z, \theta) = Bk_{-1}(\theta) + C\theta + D$. We use the fact that the initial endowments are an affine function of the skill level, $k_{-1}(\theta) = \eta_1 + \eta_2\theta$. Then:

$$\frac{\partial(J_{\varphi}(\theta) - J_{\hat{\varphi}}(\theta))}{\partial\theta} \quad = \quad \frac{1}{1-\beta}\left[\frac{\text{"constant"} + \Big[[B\eta_2 + C][\hat{B}\hat{\eta_2} + \hat{C}]\theta\Big] - \Big[[\hat{B}\hat{\eta_2} + \hat{C}][B\eta_2 + C]\theta\Big]}{[Bk_{-1}(\theta) + C\theta + D][\hat{B}k_{-1}(\theta) + \hat{C}\theta + \hat{D}]}\right]$$

Therefore the sign of the derivative does not depend on θ $\square$

There is an obvious abuse of notation in Proposition 1, since the set of implementable allocations are different depending on the type of heterogeneity in the economy.

Next we highlight the key factors behind the proof of Proposition 1. First, as mentioned before, because homothetic preferences, interior individual allocations of consumption and leisure are proportional to the counterpart aggregates. Therefore when comparing allocations Z and $\hat{Z}$, what is key is the ratio of the proportionality factors $\varphi(\theta)/\hat{\varphi}(\theta)$. Moreover, under the full insurance assumption, the proportionality factors are constant over time and are given by the value of the after tax total wealth that individuals would have if they would sell the full amount of labor to the firms. Under the assumption on the affine tax schedule for labor income, the after tax human wealth is linear in the productivity type. If there is no initial wealth inequality the function $\varphi(\theta)/\hat{\varphi}(\theta)$ is monotone in the productivity type, and therefore if the median type θ^m prefers Z to $\hat{Z}$ then at least half of the remaining types will also agree on the ordering over these two allocations. In the case of initial wealth heterogeneity, one way to ensure that the result holds is to assume that initial endowments are an affine function of the skills.

It is important to emphasize that the role of the affine tax schedule assumption is central to the above construction. In particular, it is key the fact that the after tax human wealth is linear in the productivity type. Certainly this would not be true for a general class of nonlinear tax schedules.

Next, we briefly discuss another environments in which the consensus result can be replicated. First, it is challenging to relax the assumption on the homotheticity of preferences. The main reason is that small perturbations on preferences would make the whole distribution of after-tax wealth in the economy to matter significantly.

It is possible to relax the complete markets assumption slightly. We can prove that the median voter result still holds if there is limited enforcement as in Kehoe and Levine (2001). Specifically, in an environment in which agents can buy one-period state contingent bonds (but cannot borrow), it can be shown that this additional constraint does not prevent the individual decisions from being a constant proportion of the aggregate allocations. Of course, the set of implementable allocations under complete markets is strictly bigger than the set under limited enforcement.

Proposition 1 also would be true in an environment in which there is no lump-sum component in the fiscal policy, but the government collects the taxes revenues in each period and redistributes it trough a public good g_t. In this case, the utility $v(g_t)$ that individuals get from g_t should enter additively in the period utility function.[7]

Given the linearity restriction imposed in the Proposition 1, we shall assume the following.

[7] The linearity condition does not mean that the initial distribution of capital is linear itself.

Assumption 1: The initial endowments are an affine and increasing function of skills among types: $k_{-1}(\theta) = \gamma_k + (K_{-1} - \gamma_k) \cdot \theta$ with $0 \leq \gamma_k \leq K_{-1}$.

Finally, without a restriction on the value of the lowest type $\underline{\theta}$, Proposition 1 establishes the consensus result only for fiscal policies that support interior equilibria. Without any such restriction, for some fiscal policies there will be aggregate allocations in which the decentralized competitive equilibrium exhibits a positive measure of agents supplying zero labor in equilibrium. Usually such aggregate allocations have the feature that the lump-sum component of the tax schedule is too large (a positive transfer), making too costly for the lowest types to work. These types will better off not working at least in some periods. For more details see Piguillem and Schneider (2007). In order to avoid considering economies in which non-interior allocations exist, we present a lemma that will be used to impose a lower bound on the value of $\underline{\theta}$.[8]

Lemma 2. *Consider any Z satisfying conditions (1)-(2) in Lemma 2 and having both aggregate labor sequence bounded away from zero and $\widetilde{W}_0(\theta, T, \tau_0)/\theta > 0$. There exists $\widehat{\theta} < 1$ such that $\varphi(Z;\theta) \leq \frac{1}{1-L(s^t)}$ for all s^t, $\theta \geq \widehat{\theta}$, and $L(s^t) \subset Z$.*

Proof: See Appendix A.

The lemma above provides a minimum value for $\underline{\theta}$ such that, even for the maximum feasible level of transfers $-\overline{T} > 0$ (in a competitive equilibrium with aggregate

[8] The analysis would be much more complicated in this case. To date, we can show only that the median voter result holds if we consider non-interior equilibrium in a one period economy.

labor bounded away from zero), the lowest type will work a positive amount in any period and state of nature. It also imposes a restriction on the variance of the distribution of skills.

By assumption 1, equation (10) states that individual labor supply is a monotone function in θ. Notice that the condition $\widetilde{W}_0(\theta, T, \tau_0)/\theta > 0$ implies that individual labor is minimum for the lowest type $\underline{\theta}$ in all periods and states.[9]

Assumption 2: $\underline{\theta} \in [\widehat{\theta}, 1)$.

7 Characterization of the Condorcet Winner in the log case

The characterization of the Condorcet winner comes from the maximization of the utility for the type θ^m given that the agent has to pick a sequence of aggregate allocations, an initial tax on capital and lump-sum transfers that can be supported as a competitive equilibrium. Lemma 1 gives us the sufficient implementability conditions that should be satisfied. Assumption 1 implies that is sufficient to check only the non-negativity constraint for the lowest type $\underline{\theta}$.

[9] One may ask what could happen in cases in which $\widetilde{W}_0(\theta, T, \tau_0)/\theta < 0$. We believe that in such cases the labor supply will be strictly positive for the highest type. In this case we can show there exist $\widetilde{\theta} > 1$ such that the upper bound constraint in Lemma 2, part 5, will never bind. Furthermore, for the type of distribution that we analyze in the next section, the median voter will indeed prefer $-T \geq 0$.

Then we shall partially characterize the solution for the following problem:

$$\mathbf{P(M)}: \quad \max_{\{C,L,K,T,\tau_0\}} \left\{ \frac{1}{1-\beta} \log \left(1 + (1-\beta) \left[\left(\widetilde{W}_0(\theta^m, T, \tau_{k0}) - \mathrm{E}(\widetilde{W}_0(\theta, T, \tau_{k0})) \right) \right.\right.\right.$$
$$\left.\left.\left. + (\theta^m - 1) \cdot UL \right] \right) + \sum_{t,s^t} \beta^t \Pr(s^t) \left[\alpha \log(C(s^t)) + (1-\alpha) \log(1 - L(s^t)) \right] \right\}$$

$$\text{s.t.} = \begin{cases} C(s^t) + K(s^t) \leq F\Big(L(s^t), K(s^{t-1}), s^t \Big) + (1-\delta) K(s^{t-1}) & \forall s^t \quad (\mathbf{RC}); \\ \frac{1}{C(s^t)} \geq \beta \sum_{s_{t+1}} \Pr(s_{t+1}|s^t) \frac{[1+(1-\overline{\tau})(F_k(s^{t+1})-\delta)]}{C(s^{t+1})}, & \forall s^t \quad (\mathbf{UB}); \\ 1 + (1-\beta) \left[\left(\widetilde{W}_0(\underline{\theta}, T, \tau_{k0}) - \mathrm{E}(\widetilde{W}_0(\theta, T, \tau_{k0})) \right) + (\underline{\theta} - 1) UL \right] \geq 0 & (\mathbf{NN}); \\ 1 = (1-\beta)[\mathrm{E}(\widetilde{W}_0(\theta, T, \tau_{k0})) + UL] & (\mathbf{MKT}); \\ \tau_{k0} \leq \overline{\tau}, \ K_{-1} \text{ given} & \end{cases}$$

If some agent were given the power to choose an implementable allocation, she would care about her own proportion of the aggregate allocations and also about the utility for the mean type. Those are the two parts in the objective function of the (median voter) problem. The "proportional part" basically depends on the difference in after tax total wealth between the mean and the median type. As the proof of Lemma 3 below shows, this share will be higher (lower) than θ^m if the tax schedule includes positive lump-sump transfers (taxes).

Remark: Since in the constraint set of P(M) the lump-sum term appears only in the equality condition MKT, clearly that constraint will not be binding. Notice also that if $T \leq 0$ then using constraint MKT one can see that constraint NN will not bind.

Next we state the first result relating the size of taxes and the distance between the median and the mean type. Basically, if the distance is zero, and if fiscal policy has only redistribution concerns (zero government spending), then taxes are zero in all periods and states.

Proposition 2. *Suppose $\theta^m = E(\theta) = 1$. Under Assumption 1, the most preferred allocation for the median type is the solution to a version of the Neoclassical Growth model in this environment with a representative agent having labor productivity equal to the unity and endowment K_{-1}. The implied taxes are given by $\tau_l(s^t) = \tau_k(s^t) = 0$ for all s^t.*

Proof. First, by assumption we have $E[\widetilde{W}_0(\theta, T, \tau_{k0})] = \widetilde{W}_0(\theta^m, T, \tau_{k0})$. But then the objective function reduces to:

$$\max_{C,L,K} \sum_{t,s^t} \beta^t \Pr(s^t) \left[\alpha \log(C(s^t)) + (1-\alpha)\log(1 - L(s^t))\right]$$

Next, by the remark above we know that MKT is not binding. Finally we show that maximizing the above objective function subject to the RC constraint only satisfies all the remaining constraints, i.e., UB, and NN. As it is well known the solution for the above problem implies no taxation. Since there is no government debt by assumption, we get that $T = 0$. This immediately implies that constraint NN is satisfied. But since $\tau_k(s^t) = \tau_l(s^t) = 0 \ \forall \ s^t$ we have that UB is not binding $\square$

The only way that the median voter can take advantage of nonzero marginal taxes is through the difference between the value of his wealth (initial wealth and the market value of labor endowment) and the mean wealth. When such a difference does not exist, marginal taxes are always zero. In the case where there is a sufficiently small process for government spending, the chosen fiscal policy will use only lump-sum taxes to finance the stream of expenditures. We refer to a sufficiently small process because otherwise the poorest individual in the economy may not afford the payment of the lump-sum tax.

Other than the result about taxes, the claim above adds a new interpretation to the Neoclassical Growth Model with homothetic preferences. That is, its solution can be thought also as the aggregate competitive equilibrium allocation that would be chosen by majority voting at time zero in an economy having heterogeneous labor skills drawn from a non-skewed distribution.

Next we turn the case where $\theta^m \neq E(\theta)$. Lemma 3 will be important later. It states that constraint NN will never bind in the solution to P(M).

Lemma 3. *If $\theta^m < 1$ then $T \leq 0$ in any solution to P(M).*

Proof: See Appendix A.

The intuition for Lemma 3 is simple. The main objective of the median voter is to achieve some degree of redistribution in her favor. This occurs only when the agent receive more resources than she pays. That is, because all the distortive taxes are linear and since $\theta^m < E(\theta)$, she always pays (receives, if taxes are negative) less than the average agent. On the other hand, given that all agents receive (pay) the same transfer, the only way for her to get some benefit from redistribution is to set the revenues from linear taxation at a positive value (pay less than the average) and the lump sum at a negative value (receiving the same as the average). When there is no government spending the difference is a net gain for the median agent.

Next we state a lemma which will be used later in Proposition 3. The result is an extension of the capital tax result in Bassetto and Benhabib (2006).

Lemma 4. ***(The Bang-Bang Property)*** *In the solution for the median voter's problem, if there exists $s^{\tilde{t}}$ such that the implied tax $\tau_k(s^{\tilde{t}}) < \overline{\tau}$ then*

$$\frac{1}{C^*(s^t)} = \beta \sum_{s_{t+1}} Pr(s_{t+1}|s^t) \frac{[1 + F_k^*(s^{t+1}) - \delta]}{C^*(s^{t+1})} \quad \forall \; s^t > s^{\tilde{t}}$$

and therefore $\tau_k(s^t) = 0$ for all $s^t > s^{\tilde{t}}$.

Proof: See the appendix.

In the lemma above, the notation $s^t > s^{\tilde{t}}$ is supposed to be understood as the histories that immediately follow s^t.

Remark: Notice that the proof above depends on the return function in P(M) being increasing in the utility of the mean type. In the general case ($\sigma \neq 1$), this may not be true, as Bassetto and Benhabib (2006) shows. When the return function is decreasing in the utility of the mean type, the proof can be adapted to show that the constraint UB is always binding.

Next we show that, when $\delta = 0$ and heterogeneity in the initial distribution of capital is sufficiently small, taxes on capital will be zero for any period $t \geq 2$.

Lemma 5. ***(Capital taxes)*** *Suppose $\theta^m < E(\theta) = 1$. Under assumption 1, if $\delta = 0$ and $(K_{-1} - \gamma_k)$ is arbitrarily close to zero, then the implied capital taxes in the solution to P(M) are:*

$$\tau_k(s^t) = \begin{cases} \overline{\tau} = 1 \ \ \textit{for}\ t = 0 \\ 0 < \tau_k(s^1) \leq \bar{\tau} \ \textit{for all}\ s^1 \\ 0 \ \textit{for all}\ s^t\ \ \textit{such that}\ t \geq 2 \end{cases}$$

Proof: See appendix.

If the value of $K_{-1} - \gamma_k$ is not small enough, then the Lemma must be modified slightly. Instead of having $\tau_k(s^t) = 0$ for all $t \geq 2$ it would be true for all $s^t > s^{\tilde{t}}$ given some finite $\tilde{t}$. This is a very well known result dating from the original work of Chamley (1986). It follows from the fact that, otherwise, the solution would exhibit $U^*_{ct} = \beta E_t[U^*_{c_{t+1}}]\ \forall t$. Since any solution should have $U^*_{ct}(s^t) < \infty\ \forall\ s^t$, and therefore $E(U^*_{ct}) < \infty$ for all t, it follows by the law of iterated expectations that $U^*_{ct} = \lim_{T\to\infty} \beta^T E_t[U^*_{c_{t+T}}]$. Since it can be shown that $\{U^*_{ct}\}_t$ is a submartingale, we can use Dobb's convergence theorem to show that this limit exists and is equal to zero. This leads to a contradiction since the constraint set is compact in the product topology. As we have pointed out in the remark right after Lemma 4, for the general Cobb-Douglas utility function the constraint UB may bind always.

Now consider the labor income tax. From the competitive equilibrium we know $F_L(s^t)(1-\tau_l(s^t)) = \frac{1-\alpha}{E(\theta)-L(s^t)} \frac{C(s^t)}{\alpha}$. Therefore $1-\tau_l(s^t) = \frac{(1-\alpha)}{E(\theta)-L(s^t)} \frac{C(s^t)}{F_L(s^t)\alpha}$. Then we have the following.

Proposition 3. *(**Labor Tax**) Suppose that $\theta^m < 1$. Then in the solution to P(M) there exists a history $s^{\hat{t}}$ such that, for all $s^t > s^{\hat{t}}$ the implied labor taxes are:*

1. $0 < \tau_l(s^t) < 1$.

2. *$\tau_l(s^t)$ depends on s^t only through $L(s^t)$:*

$$\tau_l(s^t) = \frac{(1-\theta^m)}{\varphi(Z^*;\theta^m)(1-L^*(s^t)) + (1-\theta^m)}$$

3. $\frac{\partial \tau_l(s^t)}{\partial(1-\theta^m)} > 0$.

Proof. The existence of a history $s^{\hat{t}}$ in which UB stops binding was justified in the previous paragraph.

Recall from Lemma 3 that (NN) is not binding. Let $\lambda(s^t)$ be the lagrange multipliers associated with (RC). Then the first order condition with respect to aggregate labor is:

$$\left[\frac{(1-\alpha)(\theta^m - E(\theta))\beta^t \Pr(s^t)}{\varphi(\theta^m)[1-L(s^t)]^2}\right] - \frac{(1-\alpha)\beta^t \Pr(s^t)}{[1-L(s^t)]} + \lambda(s^t) F_L(s^t) = 0 \quad \text{for } t \geq 1 \tag{12}$$

The implied tax on labor is given by:

$$1-\tau_l(s^t) = \left[\left(\frac{(1-\theta^m)}{\varphi(\theta^m)}\right)\frac{1}{1-L(s^t)} + 1\right]^{-1} \tag{13}$$

Let $H = \frac{(1-\theta^m)}{\varphi(\theta^m)} > 0$. Then (13) can be rewritten as:

$$0 < \tau_l(s^t) = \frac{H}{1 - L(s^t) + H} < 1 \quad \text{for } t \geq 1$$

or

$$\tau_l(s^t) = \frac{1}{\left[\frac{1}{1-\theta^m} - (1-\beta)\left(\frac{\alpha R_0(K_{-1}-\gamma_k)}{C_0} + UL\right)\right][1 - L(s^t)] + 1}$$

It is straightforward to check that $\frac{\partial \tau_l(s^t)}{\partial(1-\theta^m)} > 0$ □

Corollary 1. *(**Extending B&B**) Suppose that heterogeneity is restricted only to the initial wealth distribution, that is, individuals are indexed by the parameter $\omega \in [\underline{\omega}, \overline{\omega}]$ distributed with p.d.f. $g(\cdot)$, the labor skill is given by $\theta(\omega) = 1 \; \forall \; \omega \in [\underline{\omega}, \overline{\omega}]$, and $K(\omega) = \gamma_k + (K_0 + \gamma_k)\omega$. Then there exists $s^{\hat{t}}$ such that the Condorcet winner has $\tau_l(s^t) = 0$ for all $s^t \geq s^{\hat{t}}$.*

Lemma 5 and Proposition 3 provide some explanation of how heterogeneity shapes the Condorcet winner when agents have log preferences. First, as the median voter's problem illustrates, her payoff depends positively on the payoff obtained by the mean type. This implies that nonzero taxes benefits the voter only to the extend to which she can manipulate her share $\varphi(\theta^m)$ through the lump-sum transfers. The mean type always prefers zero taxes, and the share $\varphi(\theta^m)$ depends on the difference in after tax wealth between the median and the mean type.

Since positive capital taxes reduce the payoff for the mean type, Lemma 5 implies that the first two periods in the economy are sufficient to obtain full benefits from positive taxes on capital if the initial heterogeneity of capital endowments is small. On the other hand, labor taxes are always positive since the heterogeneity in labor income never disappears.

Also we find a smoothing effect on labor taxes similar to Werning (2006). Since distortions decrease the utility for the mean type, concavity implies that labor taxes should be higher in states in which aggregate labor is higher.

In Appendix A, the results about taxes in the general case ($\sigma \neq 1$) are extended. As Bassetto and Benhabib (2006) point out, depending on the magnitude of σ, capital taxes may be always at the upper bound. [10]. Taking this into consideration, we find a condition on the size of σ such that the capital taxes eventually go zero. This helps to characterize taxes on labor.

The results presented in the appendix are very close to the results in the last section with slight modifications. Instead of results for every $t \geq 2$ we have results for all $t > \hat{t}$, for some $\hat{t} \geq 2$. In addition, the results statements are weaker in the sense that they depend on σ being smaller than $(1-\theta^m)^{-1}$. The main role of the condition is to make sure that the objective function is increasing in aggregate consumption and decreasing in aggregate labor, as Lemma 9 shows.

The results are summarized as follows. Provided that the inequality in skills is not too large or, alternatively, $\sigma \leq (1-\theta^m)^{-1}$, capital income taxes will eventually be zero. Labor income taxes are always positive, increasing in inequality and state

[10] We omit this proof in the environment considered here because it is an extension of the reasoning presented in Bassetto and Benhabib (2006)

dependent. For all histories after which the upper bound constraint on capital taxes is not binding, labor taxes are given by

$$\tau_l(s^t) = \frac{(1-\theta^m)}{(\theta^m - 1)[(1-\alpha)(1-\sigma) - 1] + [1 - L(s^t)][(1-\sigma) + \sigma\chi(Z, \theta^m)]}$$

where $\chi(Z, \theta^m)$ resembles the proportionality factor $\varphi(Z; \theta)$ in the log case.

8 Stochastic Labor Skills

For simplicity we consider an economy without capital accumulation. Types are fixed but labor skills evolve stochastically over time. We work with log preferences, but what follows below holds for the more general Cobb-Douglas class of utility functions.

The resource constraint is given by $C(s^t)+g(s^t) \leq A(s^t)L(s^t)$. As before, at time zero there is no aggregate uncertainty. Productivity types are initially distributed according to a skewed distribution $F(\cdot)$ on $\Theta = [\underline{\theta}, \overline{\theta}]$, with mean $E(\theta) = 1$.

In $t = 1$ and later, skills of each type i evolve stochastically, and are correlated with the aggregate state . For each history s^t, skills are given by: $\theta^i(s^t) = \gamma(s^t) + \rho(s^t)\theta_0^i$, with $\rho(s^t) > 0$ for all s^t. This function allows us changes in the distribution of skills while the ranking remains constant. In addition, $\rho(s^t)$ and $\gamma(s^t)$ may be chosen such that the changes are a mean preserving spread of any particular state s_t. For example, let $\underline{\eta}_{B,N} = \frac{\underline{\theta}_B}{\underline{\theta}_N}$, where $\underline{\theta}_s$ is the productivity of the least skilled type in state s. Similarly define $\overline{\eta}_{B,N} = \frac{\overline{\theta}_B}{\overline{\theta}_N}$. Then it can be shown that if

$$\rho(B) = \frac{\overline{\theta}_N \overline{\eta}_{B,N} - \underline{\theta}_N \underline{\eta}_{B,N}}{\overline{\theta}_N \overline{\eta}_{B,N}(1 - \underline{\theta}_N) - \underline{\theta}_N \underline{\eta}_{B,N}(1 - \overline{\theta}_N)}$$

and

$$\gamma(B) = \frac{\overline{\theta}_N \underline{\theta}_N (\overline{\eta}_{B,N} - \underline{\eta}_{B,N})}{\overline{\theta}_N \overline{\eta}_{B,N}(1 - \underline{\theta}_N) - \underline{\theta}_N \underline{\eta}_{B,N}(1 - \overline{\theta}_N)}$$

then state B is a mean preserving spread of state N.

Under complete markets, as before, individual allocations are proportional to aggregate allocations. In this economy the constant of proportionality is given by:

$$\varphi(Z; \theta_0^i) = (1 - \beta) \left[-T \frac{\alpha}{C_0} + \sum_{s^t} \theta_t^i(s^t) UL(s^t) \right] \tag{14}$$

where $UL(s^t) \equiv \frac{(1-\alpha)\beta^t Pr(s^t)}{E_{s_t}(\theta) - L(s^t)}$.

Again, we can find the set of implementable allocations and let the agents vote at time zero. Let $\widetilde{\Xi}$ be the set of implementable allocations in this environment, and $S_{Z,\hat{Z}}$ be the set of types that weakly prefer Z to $\hat{Z}$. Then the following id true:

Claim: Let θ_0^m be the median of the initial types. Let $Z \in \Xi$ be the best allocation for the median type, and consider some arbitrary $\hat{Z} \in \widetilde{\Xi}$. Then either $[\underline{\theta}, \theta_0^m] \subseteq S_{Z,\hat{Z}}$ or $[\theta_0^m, \overline{\theta}] \subseteq S_{Z,\hat{Z}}$.

The reasoning of the claim above is the following. We can expresses the payoff for type i as the sum of the proportional part J_{φ^i} plus the aggregate part as in (11). As in the proof of proposition 1, we can calculate

$$\frac{\partial(J_{\varphi^i}(\theta^i) - J_{\hat{\varphi}^i}(\theta^i))}{\partial\theta^i} = \frac{\alpha(1-\beta)}{\hat{\varphi}(\theta_0^i)\varphi(\theta_0^i)} \sum_{t,s^t} \rho(s^t) \left(\frac{\hat{T}}{\hat{C}_0} UL(s^t) - \frac{T}{C_0} \hat{UL}(s^t) \right)$$

Clearly, the sign of the derivative above is independent of the type.

The best marginal tax on labor income for the median type is given by:

$$\tau_l(s^t) = \frac{(1-\theta_0^m)\rho(s^t)}{\varphi(\theta_0^m)(1-L(s^t)) + (1-\theta_0^m)\rho(s^t)} \tag{15}$$

Since the ranking among types is preserved, the larger the distance between the median and the mean type, the higher the labor tax on that state for a given aggregate labor quantity. The final effect on taxes is ambiguous. As equation (15) shows, the result depends on two factors. First, there is a tax smoothing effect: the larger the aggregate labor allocation, other things constant,the higher the tax. This is closely related to concavity and the fact that the median's utility depend on the utility of the mean type. The other effect is related to how the skills' distribution changes over the business cycles (its correlation with aggregate shocks). An increase in the distance between the mean and the median agent increases the gains of redistributive policies for the median voter, and therefore call for higher taxes. The opposite is true if the distance between the median and the mean decreases.

Thus, if inequality and employment are positively correlated, the effects reinforce each other and labor taxes are unambiguously higher. However, if inequality rises in periods of low employment (inequality and employment are negatively correlated) both effects act in opposite directions, turning the sign of the correlation between employment and labor taxation ambiguous. This is showed next by the use of a numerical example.

8.1 A Numerical Exercise

Consider an economy without capital accumulation and where skills evolve stochastically as in the previous section. Assume that there are only two possible states in the economy, $S = (High, Low)$. The stochastic process for the states is i.i.d (allowing for persistence will not affect the qualitative results), with $\pi_H = 0.6$ and $\pi_L = 0.4$. The initial state is $s_0 = H$

Technology is linear, $Y(s^t, L(s^t)) = A(s^t)L(s^t)$, with the aggregate productivity parameter being $A_L = 1.25$ and $A_H = 0.95$. Government consumption in each state takes on the values $G_H = G_L = 0.08$, which makes government consumption being about 17% of output. Preferences are logarithmic ($\sigma = 1$). We also set $\alpha = 0.3$ and $\beta = 0.95$.

The initial distribution is skewed, with the mean normalized to one and $\theta^m = 0.9$. In addition, we assume that in the high state the distribution of skills is always the same as the distribution at the initial period ($\rho(H) = 1$ and $\gamma(H) = 0$). In the low state, the distribution of skills is a mean preserving spread of the distribution at the initial state ($\rho(L) > 1$ and $\gamma(L) < 0$ with $\rho(L) + \gamma(L) = 1$).

In order to analyze the impact of the changes on the distribution of skills on the cyclical properties of the fiscal policy, we consider two economies: in Economy 1 the distribution of skills has low variability. The economy parameterized by $\rho_1(L) = 1.02$

and $\gamma_1(L) = -0.02$. In Economy 2, the skills distribution is more volatile than Economy 1. We achieve this by setting $\rho_2(L) = 1.05$ and $\gamma_2(L) = -0.05$. Figure 1 shows for the two economies the distribution at time zero and the distributions when aggregate state is $s = low$.

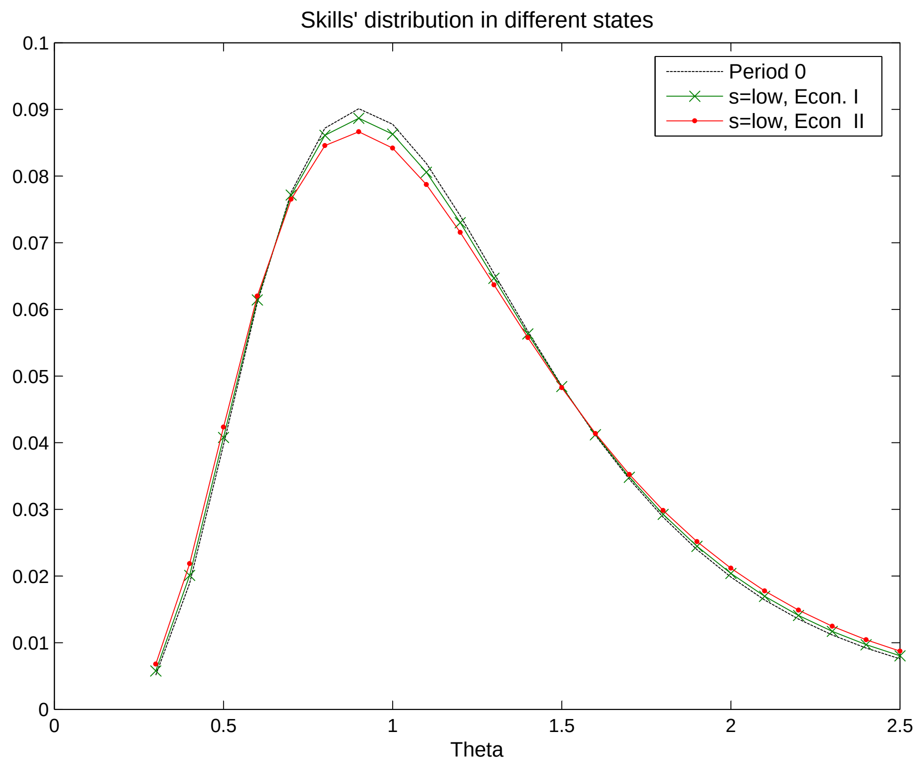

Figure 1: .

Next we show the calculated taxes for each state.

State	Economy I		Economy II	
	τ	L	τ	L
High	0.279	0.30	0.28	0.30
Low	0.278	0.285	0.285	0.28

Table 1: Aggregate labor and marginal labor tax in each state.

Labor taxes will be either procyclical or counter-cyclical, depending on how the distribution of heterogeneity changes in the low state. In Economy I, the smoothing effect predominates: higher aggregate labor implies higher taxes. In Economy II, the larger distance between the median and the mean type causes the labor tax be higher for states in which aggregate labor is lower.

The example illustrates the potential ability of the model to explain differences in the business cycles properties of labor taxes. If rich countries have sufficiently smaller dispersion in the skill distribution during bad times than poor ones, then rich countries will exhibit counter-cyclical labor taxes, while poor countries will have pro-cyclical labor taxes.

9 Quantitative Results

Several papers like Eckstein and Nagypal (2004), Heathcote, Storesletten, and Violante (2008), among others, have reported the increasing trend in labor income inequality in the U.S. in the last decades. Regarding labor income taxation, Mc-

Daniel (2007) constructs average taxes for the U.S. (and OECD countries) for the period 1950-2003 and it finds an increasing trend.

In this section we show two quantitative results of the model. First, through equation (**??**), we calculate both a lower and an upper bound on how much of the increase in labor taxes observed in the data can be accounted by the model. Second, we numerically solve the median voter's problem using a simple calibrated version of the model with stochastic labor skills. Then we compare the correlation between labor taxes and labor allocations (and GDP growth) from the model with the data.

Since the model does not have consumption taxes, we follow the same methodology as in Ohanian, Raffo, and Rogerson (2006). We compare the labor taxes from the numerical solution with $1 - \frac{1-\tau_{lt}}{1+\tau_{ct}}$ from the data. Figure 2 shows the trends in the data.

9.1 Data and Calibration

We take the average taxes on both labor and consumption for the US economy in the period 1950-2003 from McDaniel (2007).

Uncertainty is described by a Markov chain with 4 states. Both TFP and labor skills shocks are assumed to take two possible values: $A(s) \in \{A_L, A_H\}$ and $\rho(s) \in \{\rho_L, \rho_H\}$. The data for the macroeconomic aggregates are from NIPA, in billions of chained 2000 dollars covering the period 1960-2006. Since in the model we normalize the endowment of time to be equal to one, we construct a new labor series as the

Figure 2:

ratio between the total average weekly hours worked from BEA and the potential number of hours (5200 times population of 16 and over).

The production function is Cobb-Douglas with capital share $\nu = 0.3$, a usual value found in the literature. The technology parameter $A(s)$ is calibrated by using GDP from NIPA and labor as the total average weekly hours worked. The skill distribution parameters ρ_H and ρ_L are calculated from the wage inequality data in Eckstein and Nagypal (2004).[11] We take mean and median wages as a proxy for mean and median individual skills respectively. One drawback is that the data refers

[11] The authors use data from the Current Population Survey covering the period between 1961 and 2002.

to weekly earnings, and therefore it does not account for the effects of cross section variation of hours worked. But since we only need data about mean and median wages, and also given that aggregate hours worked in US has been quite stable over the last decades, we think this issue is not very critical for our purposes.

We calibrate the transition matrix over the possible four states by filtering both the TFP and the ratio median to mean wages. Then we calculate the probabilities using the frequency of the states observed in the data. The matrix below show the calculated probabilities over S=$\{s1 = (A_H, \rho_H), s2 = (A_H, \rho_L), s3 = (A_L, \rho_H), s4 = (A_L, \rho_L)\}$.

$$\begin{pmatrix} 0.438 & 0.125 & 0.375 & 0.062 \\ 0.286 & 0.286 & 0.142 & 0.286 \\ 0.455 & 0.0 & 0.091 & 0.454 \\ 0.20 & 0.20 & 0.30 & 0.30 \end{pmatrix}$$

Transition matrix.

We consider $A_H = 1 + \varepsilon_H$ and $A_L = 1 - \varepsilon_L$. We choose ε_H and ε_L such that the unconditional mean is equal to one and the process matches the variance of GDP growth in the data. In this way we set $\varepsilon_H = 0.004$ and $\varepsilon_L = 0.0064$.

The depreciation rate is set to the usual value of 0.06. We use log preferences. The parameter α is set to match, on average during the period considered, the first order conditions in the median voter's problem. Using this criterion, we find

Parameter	Description	Value
(α, σ)	preference parameters	(0.38,1)
β	intertemporal discounting	0.96
ν	capital share	0.3
δ	depreciation	0.06
(A_H, A_L)	TFP shocks	(1.004, 0.9936)
θ^m	median skill	0.79
(ρ_H, ρ_L)	skill parameters	(1.054, 0.945)

Table 2: Summary of the calibrated parameters.

$\alpha = 0.38$.[12]

Taking into account the average ratio median skill to mean skill equal to 0.79, we calibrate ρ_H and ρ_L such that we have the unconditional mean equal to one and the variance of the ratio median to mean wages matches the data. The values for ρ_H, ρ_L and the remaining parameters are summarized in Table 2 below.

We do not consider initial heterogeneity in wealth. Such strong restriction on the initial wealth distribution is imposed in order to avoid additional complications

[12]Since the first order conditions contain the share of the median type coming from the solution of his problem, the calibration for α is done in two steps. In the first step, we guess a value for the share and then calculate the average α. In the last step, using the calculate value for α, we check if the resulting share is the one that was guessed in the previous step.

related to the inequality constraints in the Euler equations. From the theory we know that if $\tau_k(s^1) < \overline{\tau}_k$, then $\tau_k(s^t) = 0$ for all $t > 1$. Since our main concern is about labor taxes, we think that initial wealth heterogeneity would add little content to the discussion at a large cost in terms of computational issues. Since the problem is not recursive, we solve it using a two-step algorithm that explores the recursive property of the Lagrangean. For more details see section A7 in the appendix 1.

The main finding of the calibrated model is a good fit of the trend in labor taxes. We assume that the conditions of Proposition 3 holds, so that labor taxes are given by (using the extension of the stochastic labor skills case):

$$\tau_l(s^t) = \frac{\rho(s^t)(1-\theta_0^m)}{\varphi(Z^*;\theta^m)(1-L(s^t)))+\rho(s^t)(1-\theta_0^m)}$$

Assuming that $\rho_t(1-\theta_0^m)$ is the actual realization of $\rho(s^t)(1-\theta_0^m)$, we can calculate both an upper and a lower bound on the process for labor taxes. These bounds come from the proof of Lemma 3 in the appendix 1: in the solution to the median voter's problem, his share is less than the unit and larger than the initial realization in skills. In order to minimize the effects of the choice of the initial period, we set $0.79 \leq \varphi(Z^*;\theta^m) < 1$, where the lower bound is given by the average value of skills in the data.

In Figure 3 we show the bounds on the process for labor taxes. Since we do not use the numerical solution of the model to calculate these boundaries, we have chosen to set the aggregate labor allocation equal to its values calculated in the data. If instead we set the aggregate labor allocation to be equal to the average

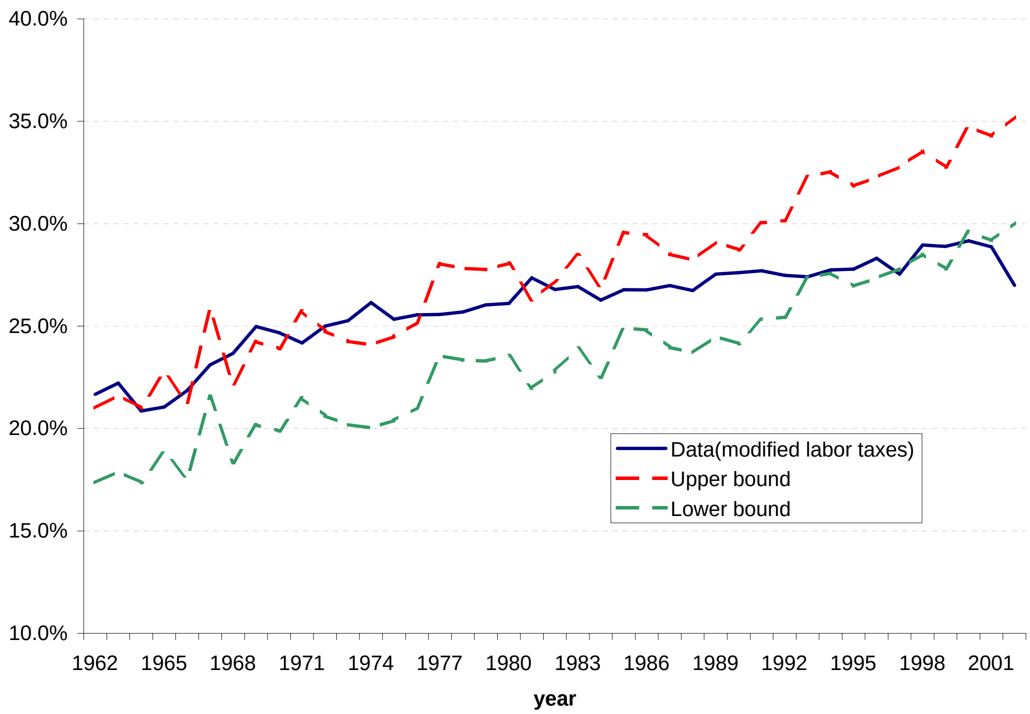

Figure 3:

value in the data, the picture would be very similar.

	Growth (%)			
	Data		Upper	Lower
	Original Taxes	Modified Taxes	Bound	Bound
1962-2001	66	33	63	68
1962-2003	43	20	67	72

Table 3: Increase in labor taxes accounted by the model.

If we consider the period 1962-2001[13], the model accounts for about two times the growth of labor taxes observed in the data. In appendix 3 we show the same

[13]This specific period does not include the significant decrease in average labor taxes between 2002 and 2003.

picture, but for the case where $\sigma = 2$.

Next we highlight some statistical properties of the calibrated economy. Moreover, we compared some properties in the data with two specifications of the calibrated model. In the first specification, labor skills are constant over time. The second specification is the one with stochastic labor skills. As one can see in Table 3 below, the model with constant skills yields almost zero variation in labor taxes, in line with the findings in Chari, Christiano, and Kehoe (1994). When we feed in the model the variation in skills to match the variance of the ratio median to mean wages, the economy matches the signal of the comovements between taxes and the relevant aggregates. The discussion in section 6 points out that the model without skills shocks should yield a correlation between labor and taxes equal to one. Since in the data the correlation between inequality of wages and TFP is negative, a priori the sign of the correlation between labor and taxes is ambiguous. The net effect in the calibrated model changes the correlation between labor and taxes to negative.

	Statistics				Correlations		
	St.Dev. ΔY	St.Dev. L	Mean Tax	St.Dev. tax	τ-ΔY	τ-ΔL	τ-L
Data	**0.02**	**0.01**	**0.26**	**0.02**	**-0.4**	**-0.1**	**- 0.7**
Model Fixed Types	**0.02**	0.002	0.26	0.001	0.7	0.2	1
Model Changing Types	**0.02**	0.006	0.26	**0.009**	-0.6	-0.2	-1

τ : Labor income tax
Y: Aggregate output
L: Aggregate labor

Table 4: Selected statistical properties of the model, $\sigma = 1$.

Finally, it turns out that the signal of the correlations presented in Table 4 are sensitive to the chosen value for σ. Table 5 shows the correlations when $\sigma = 2$ and the appropriate change in the values for (ρ_H, ρ_L) is done.

	Statistics				Correlations		
	St.Dev. ΔY	St.Dev. L	Mean Tax	St.Dev. tax	τ-ΔY	τ-ΔL	τ-L
Data	**0.02**	**0.01**	**0.26**	**0.02**	**-0.4**	**-0.1**	**- 0.7**
Model Changing Types	**0.02**	0.005	0.26	**0.01**	0.7	0.2	0.9

τ : Labor income tax
Y: Aggregate output
L: Aggregate labor

Table 5: Selected statistical properties of the model, $\sigma = 2$.

10 Conclusion

In this paper we show how heterogeneity shapes redistributive fiscal policy when individuals have balanced growth preferences and are heterogeneous with respect to both labor skills and initial wealth.

We show that the best tax sequence for the type with the median labor productivity cannot be defeated by any other policy. If only one dimensional heterogeneity is considered, i.e., either labor productivity or initial capital heterogeneity, no additional assumption regarding the distribution of types is needed. When both types of heterogeneity are taken into account simultaneously, a linear restriction about the initial wealth is required.

Regarding the characterization of the most preferred allocation by the median type, we show that if her skill is less than the mean, labor taxes are state dependent and always positive. Using a partial derivative argument at the solution, we show that labor taxes are increasing in the distance between the mean and median labor productivity. The results regarding the capital taxes are the same as in Bassetto and Benhabib (2006): taxes are always either zero or at the upper bound.

Through most of the paper we assume that skills are constant, which implies that inequality is independent of the economy's aggregate state. When skills evolve stochastically over time, but preserve the ranking among agents, a temporary increase in inequality could imply either higher or lower labor taxes, depending on the sign and level of the correlation between inequality and aggregate labor. In an economy without capital accumulation, we present a numerical example where both cases can occur.

A numerical solution for the model without capital accumulation is provided. We found that labor taxes inherit too much the changes in inequality measure by the distance between mean and median of wages in US.

The findings presented here may be useful for economies in which voting occurs sequentially over time. Also the strategy of the proof for the median voter result may be used in economies in which agents decide over objects other than taxes, but preferences are homothetic.

Part III

Sequential Voting and Time Consistent Redistribution[14]

11 Introduction

This Chapter analyzes the role of reputation in influencing political outcomes about fiscal policy. We consider a class of dynamic economies populated by agents that are heterogeneous with respect to initial wealth and vote sequentially over possible fiscal policies. This fiscal policy is restricted to redistributive linear taxes on accumulated wealth and labor income. In each period society decides the current fiscal policy according to a variation of the majority voting concept in Bernheim and Slavov (2009). The basic feature of this equilibrium concept is that current choices of both taxes and individual decisions imply different continuation policies. The first result is the existence of a Dynamic Condorcet Winner which coincides with the preferences of the median voter, the agent holding the median wealth at time zero. When agents have balanced growth preferences one can replace the majority voting criterion by the following condition about the equilibrium candidate: after any history of taxes, the median voter should prefer the outcome path generated by the candidate equilibrium to any other possible path, holding individual's decisions fixed. The second

[14] This chapter is coauthored with **Anderson Schneider**

result is a characterization of equilibrium outcomes when preferences are logarithm, production function is Cobb-Douglas and there is full depreciation of capital. The most preferred competitive equilibrium for the median type with commitment can be sustained as an equilibrium outcome for high enough discounting.

12 Economy

There is a unit measure of agents in some interval I$\subseteq \mathbb{R}$. Each agent is indexed by the initial capital stock $k_0^i \in [\overline{k_0}, \underline{k_0}]$. The distribution of initial capital is represented by the c.d.f. $F(\cdot)$ with mean K_0 and median denoted by $k_0^m \leq K_0$, unique by assumption. Let $\theta^i \equiv \frac{k_0^i}{K_0}$ represents the proportion of type i in the initial wealth. Also consider $\Theta = \{\theta^i | \exists\ k_0^i \in [\overline{k_0}, \underline{k_0}] \text{ with } \theta^i \equiv \frac{k_0^i}{K_0}\}$.

Time is discrete and the economy lasts forever. At each point in time there are three goods in the economy: consumption good, capital and labor. The output at time t is produced by competitive firms using a Cobb-Douglas production function. The aggregate resource constraint is given by:

$$C_t + K_{t+1} \leq F(L_t, K_t) + (1-\delta)K_t \tag{16}$$

Each agent has an endowment of one unit of time in each period. If agent type θ^i consumes the stream $\{c_t^i, 1 - l_t^i\}_{t=0}^{\infty}$ of consumption and leisure, then its total discounted utility is given by:

$$(1-\beta)\sum_{t=0}^{\infty}\beta^t u(c_t^i, 1-l_t^i)$$

where we assume balanced growth preferences, i.e., $u(c, 1-l) \equiv \frac{\left[c^{\alpha}(1-l)^{1-\alpha}\right]^{1-\sigma}}{1-\sigma}$ if $\sigma \neq 1$, or $\alpha \log(c) + (1-\alpha)\log(1-l)$ if $\sigma = 1$.

Firms rent labor and capital from households paying rental prices w_t and r_t respectively. In each period the government levies marginal taxes on both labor and capital income that are used for redistribution. The sequence of budget constraints is given by:

$$c_t^i + a_{t+1}^i \leq r_t(1-\tau_{kt})a_t^i + +w_t(1-\tau_{lt})l_t^i + T_t$$

where $T_t = F_{kt}K_t\tau_{kt} + F_{lt}L_t\tau_{lt}$ and $a_t^i \in A^i \equiv [\underline{a}^i, \overline{a}^i]$ is the bond holdings of agent i. The bounds $[\underline{a}, \overline{a}]$ are chosen to be large enough such that the constraint is never binding when $r_t(1-\tau_{kt}) > 0$.

We use the following notation to denote the distribution of wealth in the economy at the beginning of time t: $\Delta\mathbf{a}_t \equiv (\mathbf{a}_t^i)|_{\{i \in I\}}$.

13 Dynamic Condorcet Winner Equilibria

Next we proceed in order to define an equilibrium. The definition is similar to Bernheim and Slavov (2009) in an economy having a state variable. It is not exactly equivalent since we do not constraint the analysis to one shot deviations only.

Let $h^t = (\tau_0, \tau_1, ..., \tau_t)$ be the histories of previous policies, with $\tau_t \in [-\underline{\tau}_l, \overline{\tau}_l) \times [0, \overline{\tau}_k]$. Let $H^{-1} = \emptyset$ and $H^t = ([-\underline{\tau}_l, \overline{\tau}_l) \times [0, \overline{\tau}_k])^t$.

Definition 2. *A* ***fiscal policy (FP) program*** *is* $\pi = \{\pi_t\}_{t \geq 0}$ *with* $\pi_t : H^{t-1} \to [-\underline{\tau}_l, \overline{\tau}_l] \times [0, \overline{\tau}_k]$ *for all* $t \geq 0$*.*

Let Π be the set of FP programs. For each h^t, let $\mathcal{A}_t^i(h^t) \equiv (c_t^i(h^t), l_t^i(h^t), a_{t+1}^i(h^t))$.

Definition 3. *An* ***allocation mapping*** *is* $\mathcal{A} \equiv (\mathcal{A}^i)|i$ *such that* $\mathcal{A}^i = \{\mathcal{A}_t^i\}_{t \geq 0}$ *with* $\mathcal{A}_t^i : H^t \to \mathbb{R}^3$ *for all* $t \geq 0$*.*

A pair $(\pi, \mathcal{A})$ generates a unique outcome path $\{\tau_t, \mathcal{A}_t\}_{t \geq 0}$ from the history $h_{-1} = \emptyset : \tau_0 = \tau_0(h_{-1})$, $(c_0^i, l_0^i, a_1^i, k_1^i) = \mathcal{A}_t^i(\tau_0(h_{-1}))$ and futures fiscal policy and allocations are generated inductively by $h^t = (h^{t-1}, \pi_t(h^{t-1}))$.

A FP program π and an allocation mapping $\mathcal{A}$ induces after history $h^{t-1} \in H^{t-1}$ a continuation $(\pi, \mathcal{A})|_{h^{t-1}}$. For all $s \geq 0$, $h^s \in H^s$:

$$\pi|_{h^{t-1}}(h^s) = \pi_{t-1+s}(h^{t-1}, h^s)$$

$$\mathcal{A}|_{h^{t-1}}(h^s, \tau) = \mathcal{A}_{t-1+s}(h^{t-1}, h^s, \tau)$$

Definition 4. *For every* $h^{t-1} \in H^{t-1}$*, the* ***continuation of a FP program*** π *is* $\pi^c(h^{t-1}) = (\pi_t(h^{t-1}), \pi_{t+1}(h^{t-1}, \pi_t(h^{t-1})), ...)$*, i.e., the continuation history induced by* π*.*

For every $h^{t-1} \in H_{t-1}$, denote $\mathcal{A}(\pi^c(h^{t-1}))$ the unique outcome path of individual allocations generated by $\pi|_{h^{t-1}}$ through the allocation mapping $\mathcal{A}$.

Given $k_0^i|_i$ and a pair $\sigma = (\pi, \mathcal{A})$, we say that the policies and allocations generated by $(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$ is a **CE at time zero** if there exists a sequence of rental prices $\{r_t, w_t\}_{t=0}^{\infty}$ and transfers $\{T_t\}_{t=0}^{\infty}$ such that:

1. Given prices, taxes and transfers generated by $\pi|_{h^{t-1}}$, the sequence of allocations generated by $(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$ solves the household's problem.
2. Rental prices solve the firm's problem.
3. Aggregate resource constraint holds.

Given $(a_1^i(h_0), k_1^i(h_0))$ from a CE starting at time zero, a CE starting at history h^t can be defined recursively. For each h^t, there exists an associated distribution of wealth $\Delta\mathbf{a}_t(h^t)$.

Finally, given a pair $\sigma = (\pi, \mathcal{A})$, a history h^{t-1} and associated distribution of wealth $\Delta\mathbf{a}_t$ let $U^i(\Delta\mathbf{a}_t, \sigma|_{h^{t-1}})$ be the utility attained by household type i under the allocation generated by $(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$:

$$U^i(\Delta\mathbf{a}_t, \sigma|_{h^{t-1}}) \equiv (1-\beta)\sum_{s=t}^{\infty} \beta^{s-t} u^i(c_s^i, 1 - l_s^i/\theta^i) \tag{17}$$

Definition 5. *Given a pair $(\pi, \mathcal{A})$ and an alternative FF program $\widehat{\pi}$, if in history h^{t-1}*

$$\mu\left\{i \middle| U^i(\Delta\mathbf{a}_t, \sigma|_{h^{t-1}}) \geq U^i(\Delta\mathbf{a}_t, \widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})\right\} \geq 1/2$$

*we say that the continuation $\pi|_{h^t}$ **defeats by majority** the continuation $\widehat{\pi}|_{h^t}$ given $\mathcal{A}|_{h^{t-1}}$ and we express $(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}}) \succeq_M (\widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$.*

Notice that the continuation (mapping) $\pi|_{h^t}$ is different from $\pi^c(h^t)$ (a linked sequence of taxes).

Definition 6. *Let π be a FP program and $\mathcal{A}$ an allocation mapping. We say that $\sigma = (\pi, \mathcal{A})$ is a dynamic Condorcet winner equilibrium (**DCWE**) if:*

(i) $\forall\ h^{t-1} \in H_{t-1}$ with respective $\Delta\mathbf{a}_t$, and for all $\hat{\tau}_0 \in [-\underline{\tau}_l, \overline{\tau}_l] \times [0, \overline{\tau}_k]$, let $(\hat{\tau}, \hat{q})$ be generated by $\sigma|_{(h^{t-1}, \hat{\tau}_0)}$. Then $\hat{q} \in \Gamma(\Delta\mathbf{a}_t, \hat{\tau})$.

(ii) $\forall\ h^{t-1} \in H_{t-1}$ with respective $\Delta\mathbf{a}_t$, and $\forall\ \widehat{\pi}|_{h^{t-1}} \in \Pi$:

$$(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}}) \succeq_M (\widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$$

Given a DCW equilibrium σ, let is **value** $\Psi(\Delta k_0, \sigma)$ de given by:

$$\Psi(\Delta k_0, \sigma) = U^i(\Delta k_0, \sigma)|_i$$

Clearly, the continuation of a DCWE equilibrium is itself a DCW equilibrium.

Given an initial distribution of wealth Δk_0, the set of equilibrium payoffs is given by:

$$V = \{\Psi(\Delta k_0, \sigma) |\ \sigma \text{ is a DCWE }\}$$

14 Competitive Equilibira Characterization

The characterization of competitive equilibria turns out to be crucial for the results in this paper. In Lemma 13 we state the standard characterization in the Ramsey literature.

Lemma 13: A sequence $\{c_t^i, a_t^i, l_t^i\}_{t\geq 0}$ with interior labor and consumption is a CE allocation if and only if:

1. evaluated at the aggregate allocations, $C_t + K_{t+1} \leq F(K_t, L_t)$ holds.

2. $(1-\bar{\tau}_k)\beta F_{k_{t+1}} u^i_{c_{t+1}} \leq u^i_{ct} \leq \beta F_{k_{t+1}} u^i_{c_{t+1}}$

3. $\sum_{t\geq 0} u^i_{ct} c^i_t = u^i_{c0}(1-\tau_{k0}) F_{k0} k^i_0 + \sum_{t\geq 0} u^i_{ct}[(1-\tau_{lt})F_{lt} l^i_t + T_t]$

4. $\frac{1-\alpha}{\alpha} \frac{L_t}{1-L_t} \frac{C_t}{(1-\gamma)F(K_t,L_t)} \in [1-\overline{\tau}_l, 1-\underline{\tau}_l]$

where $T_t = \tau_{kt} F_{kt} K_t + \tau_{lt} F_{lt} L_t$

Proof: In Appendix B.

In the next Lemma we show that the each agent consumes a fixed proportion of the aggregates. The proportion for each individual is given by the (after-tax) present value of his endowments.

Lemma 14: The necessary and sufficient conditions (1)-(4) in Lemma 13 are equivalent to the existence of a function $\varphi^i : \Theta \to \mathbb{R}_+$ such that:

1' $C_t + K_{t+1} = F(K_t, L_t)$ with $[c_t^i, \theta^i - l_t^i] = \varphi^i [C_t, 1 - L_t]$

2' $(1 - \bar{\tau}_k)\beta F_{k_{t+1}} u_{c_{t+1}} \leq u_{ct} \leq \beta F_{k_{t+1}} u_{c_{t+1}}$

3' $\varphi^i = (1-\beta)\left[\frac{\alpha}{C_0}(1 - \tau_{k0})F_{k0}k_0^i + \sum_{t\geq 0} \frac{\alpha\beta^t}{C_t}((1 - \tau_{lt})F_{lt} \cdot 1 + T_t)\right]$ and $\int_I \varphi^i dF(i) = 1$

4' $\frac{1-\alpha}{\alpha}\frac{L_t}{1-L_t}\frac{C_t}{(1-\gamma)F_{lt}^{\xi}} \in [1 - \overline{\tau}_l, 1 - \underline{\tau}_l]$

where $T_t = \tau_{kt}F_{kt}K_t + \tau_{lt}F_{lt}L_t$

Proof: In Appendix B.

Because preferences are homothetic, Lemma 14 implies that, given taxes, two economies having different distributions of initial wealth but with the same mean will have the same aggregate outcomes in equilibrium. Clearly, the distribution of φ in the economy will indeed depend on the distribution of the initial endowments. Also notice that the distortions generated by marginal taxes are enclosed in the aggregates that determine the function φ. The next Lemma shows that in any point in time the wealth of individual i is an affine function of his initial share θ^i.

Lemma 15: In any competitive equilibrium, $a_t^i = K_t - \mathcal{S}_t + \mathcal{S}_t\theta^i$ for some $\mathcal{S}_t$.

Proof: The necessary and sufficient conditions in Lemma 14 generate the following wealth levels:

$$\frac{\alpha\beta^t}{C_t}(1-\tau_{kt})r_t a_t^i = \frac{\beta^t}{1-\beta}\varphi^i - \sum_{s=t}^{\infty}\beta^s\frac{\alpha}{C_s}[(1-\tau_{ls})F_{ls}+T_s]$$

Since $k_0^i = \theta^i K_0$, φ^i is affine in θ^i and therefore a_t^i is affine as well . Because $\int_\Theta a_t^i dF(i) = K_t$, we have that $a_t^i = K_t - \mathcal{S}_t + \mathcal{S}_t\theta^i$ for some $\mathcal{S}_t$ $\square$

Later we will show how $\mathcal{S}_t$ evolves over time.

For the next Lemma, it is worth to explicitly express φ^i as a function of the aggregates and the type itself. Given a interior CE, let $Z_t \equiv \{C_t, K_t, L_t\}$. Then we can express the individual share for type i as $\varphi(\{Z_t,\tau_t\}_{t\geq 0}, k_0^i)$.

Define:

$$\varphi(\{Z_t,\tau_t\}_{t\geq n}, a_n^i) \equiv (1-\beta)\left[\frac{\alpha}{C_n}(1-\tau_{kn})F_{kn}a_n^i + \sum_{t\geq n}\frac{\alpha\beta^{t-n}}{C_n}((1-\tau_{ln})F_{ln}+T_n)\right]$$

Since competitive equilibria are recursive, in the sense that a continuation of a equilibrium is itself a competitive equilibrium, the following result it is not surprising:

<u>Lemma 16</u>: Given $\tau = \{\tau_t\}_{t=0}^{\infty}$, let $\varepsilon^* = \{\{c_t^{i*}, l_t^{i*}, a_t^{i*}\}_i, C_t^*, L_t^*, K_t^*)\}_{t=0}^{\infty}$ be an interior CE allocation. Then:

$$\varphi(\{Z_t^*,\tau_t\}_{t\geq 0}, k_0^i) = \varphi(\{Z_s^*,\tau_s\}_{s\geq n}, a_n^{i*})$$

Proof: In Appendix B.

Lemma 16 states that the continuation of a CE yields the same share for type i when the economy starts at any time $t > 0$ with initial distribution of wealth $\Delta \mathbf{a}_t$. We use Lemma 4 to find a law of motion for $\mathcal{S}_t$. Since the share of each type at time zero is equal to the share at t evaluated at a_t^i, we have:

$$\frac{\alpha}{C_0}(1-\tau_{k0})r_0 K_0(\theta^i - 1) = \frac{\alpha}{C_t}(1-\tau_{kt})r_t \mathcal{S}_t(\theta^i - 1)$$

Notice that the "slope" of wealth distribution at time zero is K_0.

Since $a_t^i = (K_t - \mathcal{S}_t) + \mathcal{S}_t\theta^i$, the share of individual i in the aggregate wealth at time t is given by:

$$\theta_t^i \equiv \frac{a_t^i}{K_t} = 1 + \frac{\mathcal{S}_t}{K_t}(\theta_0^i - 1)$$

Therefore, under the assumptions on primitives, the value of the median share potentially changes over time, but the identity of the individual holding the median share is constant over time.

15 Characterization of DCW Equilibria

We start this section by showing a version of the consensus result over sequences of taxes presented in Chapter II. Roughly, the result says that, as long as heterogeneity is one-dimensional and the individual shares are an affine function of that individual characteristic, if the median voter prefers a sequence of taxes $\{\tau_s\}_{s=\hat{t}}^{\infty}$ to $\{\hat{\tau}_s\}_{s=\hat{t}}^{\infty}$, then at least 50% percent of the population also prefers $\{\tau_s\}_{s=\hat{t}}^{\infty}$ to $\{\hat{\tau}_s\}_{s=\hat{t}}^{\infty}$.

More precisely, suppose that the economy starts at date t with $\Delta \mathbf{a}_t$ given. Consider also the endowment a_t^i being an affine function of the characteristic θ_0^i. Under the assumptions on primitives, the previous characterization of competitive equilibria shows that each individual share is affine in the characteristic θ_0^i as well. Using the notation of the previous section, given a sequence of taxes, let $Z_t \equiv \{C_t, K_t, L_t\}$ be the outcome aggregate allocations at time t in the competitive equilibrium. Then we can express the individual share for type i as $\varphi(\{Z_s, \tau_s\}_{s\geq t}, a_t^i)$. Let $V^i(\{Z_s, \tau_s\}_{s\geq t}, a_t^i)$ be the present discount value of utility for type i in a competitive equilibrium starting at time t, characterized by $\{Z_s, \tau_s\}_{s\geq t}$, when he has initial endowment given by a_t^i. The difference $V^i(\{Z_s, \tau_s\}_{s\geq t}, a_t^i) - V^i(\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}, a_t^i)$ can be expressed as:

$$V^i(\{Z_s, \tau_s\}_{s\geq t}, a_t^i) - V^i(\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}, a_t^i) = \log\left(\frac{\varphi(\{Z_s, \tau_s\}_{s\geq t}, a_t^i)}{\varphi(\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}, a_t^i)}\right) + \\ + (1-\beta)\sum_{s=t}^{\infty}\beta^{s-t}\left[\alpha \log\left(\frac{C_t}{\hat{C}_t}\right) + (1-\alpha)\log\left(\frac{1-L_t}{1-\hat{L}_t}\right)\right] \quad (18)$$

Agents in this economy evaluate the difference above by comparing the difference in the (log of the) individual shares to the second term, which is not individual specific. The next proposition shows that, for every $\{Z_s, \tau_s\}_{s\geq t}$ and $\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}$, the following single-crossing property holds:

Proposition 5: Let Ξ the space of sequences $\{Z_s, \tau_s\}_{s\geq t}$ that satisfies the conditions in Lemma 14. Suppose that a_t^i is an affine function of the individual specific characteristic θ_0^i. Then, given any $\{Z_s, \tau_s\}_{s\geq t}$, $\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t} \in \Xi$, the functions $\log\left(\varphi(\{Z_s, \tau_s\}_{s\geq t}, a_t^i)\right)$ and $\log\left(\varphi(\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}, a_t^i)\right)$ cross exactly once at $a_t^i = K_t$, where:

$$\varphi(\{Z_s, \tau_s\}_{s\geq t}, a_t^i) = 1 + (1-\beta)\frac{\alpha}{C_t}(1-\tau_{_{kt}})r_t a_t^i$$

Proof: See proof in Chapter II.

Using Proposition 5 and Lemma 15, we have the following characterization of dynamic Condorcet winner equilibria.

Proposition 6: Take a pair $(\pi, \mathcal{A})$ consisting of a allocation mapping and a FP program. Then $(\pi, \mathcal{A})$ is a DCW if and only if:

1. $\forall\ h^{t-1} \in H^{t-1}$ with respective $\Delta\mathbf{a}_t$, and for all $\hat{\tau}_0 \in [-\underline{\tau}_l, \overline{\tau}_l] \times [0, \overline{\tau}_k]$, let $(\hat{\tau}, \hat{q})$ be generated by $\sigma|_{(h^{t-1}, \hat{\tau}_0)}$. Then $\hat{q} \in \Gamma(\Delta\mathbf{a}_t, \tilde{\tau})$.

2. $\forall\ h^{t-1} \in H^{t-1}$, and $\forall\ \widehat{\pi}|_{h^{t-1}} \in \Pi$:

$$U^m(\Delta\mathbf{a}_t, \sigma|_{h^{t-1}}) \geq U^m(\Delta\mathbf{a}_t, \widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$$

Proof: (If) Consider some arbitrary history h^{t-1}. The outcome path of $\sigma|_{h^{t-1}}$ is a competitive equilibrium and, by Lemma 3, it yields a distribution of wealth $\Delta\mathbf{a}_t$ which is affine in θ_0^i. Given $\Delta\mathbf{a}_t$, then by the single crossing property , if $U^m(\Delta\mathbf{a}_t, \sigma|_{h^{t-1}}) \geq U^m(\Delta\mathbf{a}_t, \widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$ we have that $U^i(\Delta\mathbf{a}_t, \sigma|_{h^{t-1}}) \geq U^i(\Delta\mathbf{a}_t, \widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$ either by all $\theta^i \leq \theta^m$ or all $\theta^i \geq \theta^m$. This implies $(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}}) \succeq_M (\widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$.

Since the history h^{t-1} was arbitrary, conditions 1 and 2 in the statement imply that $(\pi, \mathcal{A})$ is a DCW.

(Only If) Let $\{Z_s, \tau_s\}_{s\geq t}$ be the outcome path of taxes and aggregates generated by $\sigma|_{h^{t-1}}$ and $\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}$ the respective outcome generated by $\widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}}$.Suppose that $(\pi, \mathcal{A})$ is a DCW and that condition 2 does not hold at some history h^t:

$$U^m(\Delta \mathbf{a}_t, \sigma|_{h^{t-1}}) < U^m(\Delta \mathbf{a}_t, \widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$$

This can be rewritten as:

$$\log(\varphi(\{Z_s, \tau_s\}_{s\geq t}, a_t^m)) - \log(\varphi(\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}, a_t^m)) > V(\{\hat{Z}_s, \hat{\tau}_s\}_{s\geq t}) - V(\{Z_s, \tau_s\}_{s\geq t})$$

By the single-crossing property, since the median type prefers one to the other, more than the majority supports the alternative path, which contradicts $(\pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}}) \succeq_M (\widehat{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$ $\square$

Proposition 6 simplifies to a great extent the task of checking the requirements for a DCW. In particular, condition 2 in the definition of a DCW can be replaced without loss of generality in this economy by the second condition in Proposition 6. Therefore the condition that the continuation $\pi|_{h^t}$ defeats by majority any other plan, given the individual decisions of agents, is replaced by a time-consistent requirement. Moreover, time-consistency is evaluated at the median's allocation on the outcome path given the distribution of wealth after any history of taxes.

As a byproduct of Proposition 6, DCW equilibria in the economy considered here is equivalent to sustainable equilibrium (as in Chari and Kehoe (1990) and Phelan and Stacchetti (2001)) in which the government sets redistributive fiscal policy and assigns welfare weight equal to one to the median type.

In the last section of the appendix we show that the aggregate allocations are bounded in any competitive equilibrium. Given discounting, this fact allow us to restrict attention to one shot deviations from the candidate equilibrium $\pi \in \Pi$.

Definition: A FP program $\widetilde{\pi}|_{h^{t-1}} \in \Pi$ is an one-shot deviation (OSD) from program $\pi|_{h^{t-1}} \in \Pi$ at $h^{t-1} \in H_{t-1}$ if:

1. $\widetilde{\pi}_t|_{h^{t-1}} \neq \pi_t|_{h^{t-1}}$; and

2. $\widetilde{\pi}_{s-1}|_{h^{t-1}}(h^s) = \pi_{s-1}|_{h^{t-1}}(h^s)$ for all h^s, $s > t$.

Definition: A one-shot deviation (OSD) $\widetilde{\pi}|_{h^{t-1}} \in \Pi$ is profitable at $h^{t-1} \in H_{t-1}$ if, by holding $\mathcal{A}|_{h^{t-1}}$ fixed, we have:

$$U^m(\Delta \mathbf{a}_t, (\widetilde{\pi}|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}}) > U^m(\Delta \mathbf{a}_t, \pi|_{h^{t-1}}, \mathcal{A}|_{h^{t-1}})$$

Corollary In a DCWE, condition 2) is equivalent to:

$\forall\ h^{t-1} \in H^{t-1},$ does not exist a profitable OSD $\widehat{\pi}|_{h^{t-1}} \in \Pi$ from $\pi|_{h^{t-1}}$ at $h^{t-1} \in H_{t-1}$

16 Existence of Equilibrium - Log case and full depreciation

We will show an equilibrium in which, for all h^{t-1}, $\pi_t(h^{t-1}) = (\overline{\tau}_{kt}, \tau_l^*)$ for some τ_l^*. Moreover, the assumptions that both utility is logarithm and full depreciation allow us to get close form solutions. In order to show this equilibrium , first we characterize CE for policies in which $\tau_{kt} = \overline{\tau}_k \ \forall t$.

<u>Lemma 17</u>: Suppose that utility is logarithm, $\delta = 1$ and $\tau_{kt} = \overline{\tau}_k \ \forall t$. The CE given $\{\tau_{kt}, \tau_{lt}\}_{t=0}^{\infty}$ and initial distribution Δk_0 is given by:

1. $C_t^i = \varphi^i C_t, 1 - l_t^i = \varphi^i (1 - L_t)$
2. $C_t = (1 - \beta\gamma(1 - \overline{\tau}_k))F(K_t, L_t)$
3. $K_{t+1} = \beta\gamma(1 - \overline{\tau}_k)F(K_t, L_t)$
4. L_t is the unique solution to $\frac{L_t}{1-L_t} = \frac{\alpha(1-\gamma)(1-\tau_{lt})}{(1-\beta\gamma(1-\overline{\tau}_k))(1-\alpha)}$
5. $\varphi^i = 1 + (1-\beta)[\frac{\alpha}{C_t}(1 - \tau_{kt})F_{kt}\mathcal{S}_t(\theta^i - 1)]$

where $\mathcal{S}_t$ evolves according $\frac{\alpha}{C_0}(1 - \tau_{k0})F_{k0}K_0(\theta^i - 1) = \frac{\alpha}{C_t}(1 - \tau_{kt})F_{kt}\mathcal{S}_t(\theta^i - 1)$.

Next, consider some history h^t with $(K_t, \mathcal{S}_t)$ given. Imagine the situation of the median type choosing her best current tax on wealth and a sequence of **constant**

taxes on labor given that future taxes on wealth are given by $\overline{\tau}_k$ (and that the continuation of the allocation plan is given by the rule in Lemma 6 above with $\tau_{ks} = \overline{\tau}_k \; \forall s > t$ and labor taxes constant).

Given that the continuation of both the FP program and the allocation, labor is fixed over time. Then the problem reduces to:

$$\max_{\tau_{kt},\tau_l} \; \log\left(1 + (1-\beta)\frac{\alpha}{C_t}(1-\tau_{kt})F_{kt}\mathcal{S}_t(\theta^m - 1)\right) + (1-\beta)\left[\sum_{s=t}^{\infty}\beta^{s-t}\alpha\log(C_s) + (1-\alpha)\log(1-Ls)\right]$$

$$\text{s.t.} \begin{cases} L_s = L = \frac{\psi}{1+\psi} \forall s \geq t, \; \psi \equiv \frac{\alpha(1-\gamma)(1-\tau_l)}{(1-\beta\gamma(1-\overline{\tau}_k))(1-\alpha)} \\ K_s = [\beta\gamma(1-\overline{\tau}_k)L^{1-\gamma}]^{1+\gamma+\gamma^2+\ldots+\gamma^{s-1}}K_t^{\gamma^s} \\ C_s = [1-\beta\gamma(1-\overline{\tau}_k)]F(K_t, L_t) \\ K_t, \, \mathcal{S}_t \text{ given} \end{cases}$$

In the problem above, the choice of τ_{kt} only affects the share for the median type. The choice of τ_l implies a choice for labor $L_s = L$ for all $s \geq t$. Using the constraint set, the return function becomes (up to a constant):

$$\log\left(1 + (1-\beta)\frac{\alpha(1-\overline{\tau}_k)\gamma}{1-\beta\gamma(1-\overline{\tau}_k)}\frac{\mathcal{S}_t}{K_t}(\theta^m - 1)\right) + \left[\alpha(1-\gamma)\log(L) + \frac{\alpha\gamma\beta(1-\gamma)}{1-\gamma\beta}\log(L) + (1-\alpha)\log(1-Ls)\right]$$

The best choice of τ_{kt} is to set wealth taxes at the upper bound. The (constant) labor tax that maximizes the return function is $\tau_l^{trigger} = \frac{-\overline{\tau}_k\gamma\beta}{1-\gamma\beta}$. Notice that $\tau_l^{trigger}$

corrects the intratemporal distortion completely when capital taxes are set in the upper bound.

Now consider a slightly different problem given history h^t $((K_t, \mathcal{S}_t)$ given): choose (τ_{kt}, τ_{lt}) given that future taxes are $(\overline{\tau}_k, \tau_l^{trigger})$ in all periods in the future.

The return function(**up to a constant**) is given by:

$$\log(\varphi^m) + (1-\beta)\left[\alpha\log(L_t^{1-\gamma}) + (1-\alpha)\log(1-L_t)\right] + \beta[\alpha\log(K_{t+1}^{\gamma})] + \beta^2[\alpha\log(K_{t+2}^{\gamma})] + \ldots$$

where

$K_{t+1} = [\beta\gamma(1-\overline{\tau}_k)K_t L_t^{1-\gamma}]$

$K_{t+2} = [\beta\gamma(1-\overline{\tau}_k)]K_{t+1}^{\gamma}L_{t+1}^{1-\gamma} = [\beta\gamma(1-\overline{\tau}_k)]^{1+\gamma}K_t^{\gamma^2}L_t^{(1-\gamma)^{\gamma}}L^{1-\gamma}$

$K_{t+2} = [\beta\gamma(1-\overline{\tau}_k)]^{1+\gamma+\gamma^3}K_t^{\gamma 3}L_t^{(1-\gamma)^{\gamma^2}}L^{(1-\gamma)^{1+\gamma}}$

where L is determine by $\tau_l^{trigger}$

Then we can rewrite the return function as:

$$\log(\varphi^m) + (1-\beta)[\alpha(1-\gamma)\log(L_t) + (1-\alpha)\log(1-L_t)] + \alpha\gamma\sum_{s\geq t+1}\beta^{s-t}[\log((L_t^{1-\gamma})^{\gamma^{s-(t+1)}})]$$

or after some little work:

$$\log(\varphi^m) + (1-\beta)[\alpha(1-\gamma)\log(L_t) + (1-\alpha)\log(1-L_t)] + \log(L_t)\alpha\gamma\left[\beta(1-\gamma) + \beta^2(1-\gamma)\gamma + \beta^3(1-\gamma)\gamma^2 + \ldots\right]$$
$$\log\left(1 + (1-\beta)\frac{\alpha(1-\overline{\tau}_k)\gamma}{1-\beta\gamma(1-\overline{\tau}_k)}\frac{\mathcal{S}_t}{K_t}(\theta^m - 1)\right) + \left[\alpha(1-\gamma)\log(L) + \frac{\alpha\gamma\beta(1-\gamma)}{1-\gamma\beta}\log(L) + (1-\alpha)\log(1-Ls)\right]$$

as before. Therefore, given future taxes and allocations driven by the "expectation" than taxes will be $(\overline{\tau}_k, \tau_l^{trigger})$ in all periods in the future, the best choice for the median type at time t is indeed $(\tau_{kt}, \tau_{lt}) = (\overline{\tau}_k, \tau_l^{trigger})$.

Lemma 18: The following $\sigma^{trigger}$ is a DCWE:

- Fiscal Policy Program $\pi^{trigger}$: $\forall\ h^{t-1} \in H^{t-1},\ \tau_{kt}(h^{t-1}) = \overline{\tau}_k$ and $\tau_{lt}(h^{t-1}) = \tau_l^{trigger} \equiv \frac{-\overline{\tau}_k \gamma\beta}{1-\gamma\beta}$.

- Allocation plan $\mathcal{A}^{trigger}$: $\forall\ h^{t-1} \in H^{t-1},$ given inherited $\Delta\mathbf{a}_t, \forall\ (\hat{\tau}_l, \hat{\tau}_k) \in [-\underline{\tau}_l, \overline{\tau}_l] \times [0, \overline{\tau}_k]$:

  1. $L_t(h^t)$ solves $\frac{L}{1-L} = \frac{\alpha(1-\gamma)(1-\hat{\tau}_l)}{(1-\beta\gamma(1-\overline{\tau}_k))(1-\alpha)}$
  2. $L_s(h^s)$ solves $\frac{L}{1-L} = \frac{\alpha(1-\gamma)(1-\tau_l^{trigger})}{(1-\beta\gamma(1-\overline{\tau}_k))(1-\alpha)}$ for all $s > t$
  3. $C_t(h^t) = [1 - \beta\gamma(1 - \overline{\tau}_k)]F(K_t(h^t), L_t(h^t))$
  4. $K_{t+1}(h^t) = \beta\gamma(1 - \overline{\tau}_k)F(K_t(h^t), L_t(h^t))$
  5. $c_t^i(h^t) = \varphi^i C_t(h^t),\ 1 - l_t^i(h^t) = \varphi^i(1 - L_t(h^t))$
  6. $\varphi^i = 1 + (1 - \beta)\frac{\alpha}{C_t(h^t)}(1 - \hat{\tau}_k)F_{kt}\mathcal{S}_t(\theta^i - 1)$

Proof: We use the OSD result.First, given future policies and allocation plan, we just have shown above that the best choice for the median type, after any history, is to set $(\tau_{kt}, \tau_{lt}) = (\overline{\tau}_k, \tau_l^{trigger})$. Therefore there is no profitable OSD evaluated at median's utility. Finally, for all $(\hat{\tau}_l, \hat{\tau}_k) \in [-\underline{\tau}_l, \overline{\tau}_l] \times [0, \overline{\tau}_k]$, the continuation of the allocation plan is a CE given taxes generated by π $\square$

Given $\sigma^{trigger}$ and history h^{t-1} with $(K_t, \mathcal{S}_t)(h^{t-1})$, denote by $V^{m,trigger}(K_t, \mathcal{S}_t)$ the value of the utility for the median type under $\sigma^{trigger}|_{h^{t-1}}$.

Proposition 7: Let $x^* = \{C_t, L_t, K_t, \mathcal{S}_t\}_{t=0}^{\infty}$. Then x^* is the outcome of a DCWE if:

1. x^* is a CE at time zero.

2. $\sum_{s=t}^{\infty}(1-\beta)\beta^{s-t}u^m(x_s) \geq V^{m,trigger}(K_t, \mathcal{S}_t)$ for all $s \geq 0$.

Proof: Suppose we have such a sequence x^*. Then we construct a equilibrium that sets x^* as the equilibrium path (taxes calculated from x^* as usual) and that calls for reversion to $\sigma^{trigger}$ if a deviation occurs in any history. Using the condition, such a sequence is the outcome of a DCWE.

17 Commitment Solution is an Outcome of a DCW

Neglecting the bounds on labor tax, under commitment the best tax sequence for the median type is given by:

$$\max_{\tau_{k0},C,L,K} \frac{1}{1-\beta}\log\left(1+\frac{\alpha}{C_0}(1-\beta)\gamma F(K_0,L_0)(\theta^m-1)\right)+\sum_{t=0}^{\infty}\beta^t[\alpha\log(C_t)+(1-\alpha)\log(1-L_t)]$$

$$\text{s.t.}\begin{cases} C_t + K_{t+1} \leq F(K_t, L_t) \\ C_{t+1} \geq \beta(1-\overline{\tau}_k)F_{k_{t+1}}C_t \;\; t \geq 0 \end{cases}$$

Let λ_t and ψ_t be the multiplier associated with the first and second constraint respectively .

Lemma 19: Let $\overline{L} = \frac{\alpha(1-\underline{\tau}_l)}{(1-\underline{\tau}_l)+(1-\alpha)\underline{\tau}_l}$. If

$$\frac{\gamma(1-\overline{\tau}_k)(1-\theta^m)}{\frac{\alpha(1-\overline{L})}{(1-\alpha)\overline{L}}+\gamma(1-\overline{\tau}_k)(1-\theta^m)(1-\alpha(1-\beta))} < \overline{\tau}_k$$

then $\psi_t = 0$ for all t in the solution to the problem above.

Proof: In Appendix B.

Using the lemma above, we have that the sequence of taxes $\{\tau_t^*\}_{t\geq 0}$ in the commitment solution is:

1. $\tau_{k0}^* = \overline{\tau}_k,\ 0 < \tau_{k1}^* =< \overline{\tau}_k,\ \tau_{ks}^* = 0\ s \geq 2$
2. $\tau_{l0}^* < 0,\ \tau_{ls}^* = 0\ s \geq 1$

The aggregates in the CE generated by $\{\tau_t^*\}_{t\geq 0}$ are:

1. $C_0 = \frac{1-\beta\gamma}{1-\beta\gamma\tau_{k1}^*}F(K_0, L_0),\ \ K_1 = \frac{\beta\gamma(1-\tau_{k1}^*)}{1-\beta\gamma\tau_{k1}^*}F(K_0, L_0)$
2. $C_s = (1-\beta\gamma)F(K_s, L_s),\ K_{s+1} = \beta\gamma F(K_s, L_s)\ \ s \geq 1$
3. L_0 solves $\frac{L_0}{1-L_0} = \frac{\alpha(1-\gamma)(1-\tau_{l0}^*)}{\frac{1-\beta\gamma}{1-\beta\gamma\tau_{k1}^*}(1-\alpha)}$
4. L_s solves $\frac{L_s}{1-L_s} = \frac{\alpha(1-\gamma)}{(1-\beta\gamma)(1-\alpha)}$ for $s \geq 1$.

Claim: In the outcome path of the commitment solution, under condition of Lemma 19, we have:

$$\frac{\mathcal{S}_1}{K_1} = \frac{(1-\overline{\tau}_k)(1-\beta\gamma\tau^*_{k1})}{1-\tau^*_{k1}}, \quad \frac{\mathcal{S}_t}{K_t} = (1-\overline{\tau}_k)(1-\beta\gamma\tau^*_{k1})$$

Proof: Along the path of any CE we have:

$$\frac{\alpha(1-\tau_{k_0})F_{k0}K_0}{C_0} = \frac{\alpha(1-\tau_{k_t})F_{kt}\mathcal{S}_t}{C_t} \;\; \forall\; t$$

Given the characterization of aggregates above we get the claim $\square$

Proposition 8: If preferences are logarithm and $\delta = 1$, there exists $\beta^* \in (0,1)$ such that the solution to the median voter's problem under commitment is the outcome of a DCW.

In order to check that the commitment solution is an outcome of a DCW, we use Proposition 7. For $t \geq 1$, $L^{trigger} = L^*_t$ and:

$$V^m(K^*_t, \mathcal{S}^*_t) - V^{m,trig}(K^*_t, \mathcal{S}^*_t) = \log\left(\frac{\varphi^m}{\varphi^{m,trig}_t}\right) + \alpha\log\left(\frac{1-\beta\gamma}{1-\beta\gamma(1-\overline{\tau}_k)}\right) - \frac{\alpha\beta\gamma}{1-\beta\gamma}\log(1-\overline{\tau}_k)$$

Also $\varphi^{m,trig}_1 > \varphi^m$ if $\beta < 1$, with equality if $\beta = 1$. The last term in the expression above, $g(\beta)$, is such that $g(0) = 0$ and $g'(\beta) > 0$. Therefore $\exists\, \widehat{\beta} < 1$ such that the difference is positive. Similar reasoning remains for the condition 2) in Proposition 7 at $t = 2$.

References

ALESINA, A., AND D. RODRIK (1994): "Distributive Politics and Economic Growth," *The Quarterly Journal of Economics.*

AZZIMONTI, M., E. DE FRANCISCO, AND P. KRUSELL (2008): "Aggregation and Aggregation ," *Journal of the European Economic Association*, 6, 381–394.

BASSETTO, M., AND J. BENHABIB (2006): "Redistribution, Taxes, and the Median voter," *Review of Economic Dynamics*, (9).

BENHABIB, J., AND A. PRZEWORSKI (2006): "The Political Economy of Redistribution under Democracy," *Economic Theory.*

BERNHEIM, B., AND S. N. SLAVOV (2009): "A Solution Concept for Majority Rule in Dynamic Settings," *Review of Economic Studies*, 76, 33–62.

CHAMLEY, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica.*

CHARI, V., L. CHRISTIANO, AND P. KEHOE (1994): "Optimality fiscal policy in a business cycle model," *Journal of Political Economy*, 102, 617–652.

CHARI, V., AND P. KEHOE (1990): "Sustainable Plans," *Journal of Political Economy*, 98, 783–802.

CHARI, V., AND P. J. KEHOE (1999): "Optimal Fiscal and Monetary Policy," *in J.B. Taylor and M. Woodford (ed.), Handbook of Macroeconomics.*

Corbae, D., P. D'Erasmo, and B. Kuruscu (2009): "Politico-economic consequences of rising wage inequality ," *Journal of Monetary Economics*, 56, 43–61.

Eckstein, Z., and E. Nagypal (2004): "The Evolution of US Earnings Inequality: 1961-2002," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28, 10–29.

Heathcote, J., K. Storesletten, and G. Violante (2008): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *NBER Working Paper*, 14052.

Kehoe, T., and D. Levine (2001): "Liquidity Constrained Markets versus Debt Constrained Markets," *Econometrica*.

Kramer, G. (1973): "On a Class of Equilibirum Conditions for Majority Rule," *Econometrica*, 41(2).

Krusell, P., and J. Rios-Rull (1999): "On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model," *The American Economic Review*.

McDaniel, C. (2007): "Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003," *Working Paper*.

Meltzer, A., and S. Richard (1981): "A Rational Theory of the Size Government," *Journal of Political Economy*.

MILGROM, P., AND I. SEGAL (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70(2).

OHANIAN, L., A. RAFFO, AND R. ROGERSON (2006): "Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004," *FRB of Kansas City Working Paper No. 06-16.*

PERSSON, T., AND G. TABELLINI (1994): "Is Inequality Harmful for Growth?," *American Economic Review*, 84, 600.

PHELAN, C., AND E. STACCHETTI (2001): "Sequential Equilibrium in Ramsey Tax Model ," *Econometrica*, 69, 1491–1518.

PIGUILLEM, F., AND A. L. SCHNEIDER (2007): "A note on redistribution, optimal fiscal policy and corner solutions," *(www.econ.umn.edu/ ~ als/Research/Taxescornersol.pdf).*

WERNING, I. (2006): "Optimal Fiscal Policy with Redistirbution," *Quarterly Journal of Economics.*

Part IV

Appendix A: proofs of Chapter II

17.1 Proof of Lemma 1:

Here we prove the sufficiency of conditions (1)-(3). First, use the function $\varphi(\cdot)$ and Z to construct the individual consumption and labor allocations.

Set after-tax prices as:

$$p(s^t) = \beta^t \Pr(s^t) U_c^{CE}(s^t) = \beta^t \Pr(s^t) E(\varphi)\alpha / C(s^t), \;\; p(s_0) = \alpha / C_0$$

$$p(s^t) w(s^t)(1 - \tau_l(s^t)) = \beta^t \Pr(s^t) E(\varphi)(1-\alpha)/(1 - L(s^t))$$

And

$$r(s^t) = F_k(s^t), w(s^t) = F_L(s^t), p(s^t) = \sum_{s_{t+1}} p(s^{t+1}) R(s^{t+1})$$

If $0 < \frac{\varphi(\theta)}{E(\varphi)} \leq \frac{\theta}{1-L(s^t)}$ for all s^t, then one can check, using the solution to the static problem, that the following necessary first order conditions are met for all s^t and $l \in [0,1]$:

$$\left[U_l(c^*(s^t;\theta), 1 - l^*(s^t;\theta)/\theta) + U_c(c^*(s^t;\theta), 1 - l^*(s^t;\theta)/\theta) \Big(w(s^t)(1 - \tau_l(s^t)) \Big) \right] [l - l^*(s^t;\theta)] \leq 0$$

The transversality condition (Tvc) $\lim_{t \to \infty} \sum_{s^t} p(s^t) k(s^t;\theta) = 0$ is satisfied because it can be shown that individual capital allocations are an affine function of the aggregate capital stock. At the equilibrium prices, the Tvc is met, since the aggregate allocations are bounded in the product topology. Finally, using 10, we can get the budget constraint back. Condition (2) in the competitive equilibrium definition is satisfied by construction. As usual, the government budget constraint can be recovered using a version of the Walras' law. Taxes on capital can be constructed in many ways, and taxes on labor are constructed using the definition of prices and $w(s^t) = F_L(s^t)$ ■

17.2 Proof of Lemma 2:

Let $\widetilde{\Xi}$ be the set of allocations $Z \equiv (\{C(s^t), L(s^t), K(s^t)\}_{t\geq 0}, T \leq 0, \tau_{k0})$ with aggregate labor allocation bounded away from zero, the resources constraint satisfied for all periods, and the Euler equation satisfied with weak inequality.

For any $Z \in \widetilde{\Xi}$, let $\underline{L}(Z) \equiv \inf\{L(s^t)\}_{t\geq 0}$ and define $\underline{\theta}(Z)$ to be the solution to:

$$\inf \quad \theta \quad \text{s.t.} \begin{cases} \theta \in [0,1] \\ \varphi(Z;\theta) < \frac{1}{1-\underline{L}(Z)} \end{cases}$$

Claim: $\underline{\theta}(Z)$ is bounded away from 1 for all $Z \in \widetilde{\Xi}$.

Proof of the claim: Because of the linearity of $\varphi(\cdot)$ in types, it follows that $\varphi(Z;\theta = 1) = 1$. $\{L(s^t)\}_{t\geq 0}$ is bounded away from zero, and therefore $\frac{1}{1-\underline{L}(Z)} \geq \frac{1}{1-\epsilon}$ for some $\epsilon > 0$. The claim follows.

Define $\widehat{\theta} \equiv \sup \{\underline{\theta}(Z) : Z \in \widetilde{\Xi}\}$. Because of the claim above, $\widehat{\theta} < 1$. Then it is straightforward to check that $\widehat{\theta}$ has the property stated in the Lemma. In particular, if $\underline{\theta}(Z)$ satisfies the second constraint in the inf problem above, then it satisfies that constraint for all $L(s^t) \subseteq Z)$ □

17.3 Proof of Lemma 3:

If the statement is not true, then in the solution to $P(M)$ we have $T^* > 0$. The value of the program $P(M)$ can be written as:

$$P(M) = \frac{1}{1-\beta}\log(\varphi(Z^*, \theta^m)) + V(Z^*)$$

where $\varphi(Z^*; \theta^m)$ and $V(Z^*)$ are given by (10) and the last part of (11) respectively, evaluated at Z^*.

Now fix $\hat{T} = 0$. For any $\hat{Z} \in \Xi$ with $\hat{T} = 0$

$$\begin{aligned}
\varphi(Z; \theta^m) &= (1-\beta)\theta^m \left[\frac{\alpha}{\hat{C}_0}\hat{R}_0\frac{(\gamma_k)}{\theta^m} + \frac{\alpha}{\hat{C}_0}\hat{R}_0(K_{-1} - \gamma_k) + \widehat{UL}\right] \\
&= \theta^m + \varepsilon(\hat{Z})
\end{aligned}$$

for some $\varepsilon(\hat{Z}) > 0$. The last line comes from the constraint (MKT), $\theta^m < 1$ and $\gamma_k > 0$.

Next, define the feasible allocation $\hat{Z} \in \Xi$ with $\hat{T} = 0$ as:

$$\hat{Z} \in \text{argmax}_{\{C,L,K,\tau_{k0}\}} V(Z)$$

$$\text{s.t.} = \begin{cases}
C(s^t) + K(s^t) \leq F\Big(L(s^t), K(s^{t-1}), s^t\Big) + (1-\delta)K(s^{t-1}) & \forall s^t \quad (\textbf{RC}); \\
\frac{1}{C(s^t)} \geq \beta \sum_{s_{t+1}} \Pr(s_{t+1}|s^t)\frac{[1+(1-\overline{\tau})(F_k(s^{t+1})-\delta)]}{C(s^{t+1})}, & \forall s^t \quad (\textbf{UB}); \\
(1-\beta)\left[\frac{\alpha}{C_0}R_0 k_{-1}(\underline{\theta}) + \underline{\theta}UL\right] \geq 0 & (\textbf{NN}); \\
\tau_{k0} \leq \overline{\tau}, \ K_{-1} \text{ given}
\end{cases}$$

Clearly, the constraint NN will never bind. Therefore the value of $\hat{Z}$ in terms of utility is given by:

$$\hat{P}(Z) = \frac{1}{1-\beta}\log(\theta^m+\varepsilon(\hat{Z}))+\left\{\max_{\{C,L,K,\tau_{k0}\}\in(RC,UB,\tau_{k0}\leq\mathcal{T})} V(Z)\right\} = \frac{1}{1-\beta}\log(\theta^m)+V(\hat{Z})$$

Then, since Z^* solves $P(M)$, we have that:

$$P(\widehat{Z})-P(M) = \frac{1}{1-\beta}\left[\log(\theta^m+\varepsilon(\hat{Z})) - \log\left((1-\beta)\left(\frac{\alpha}{C_0^*}(R_0^*(\gamma_k+(K_{-1}-\gamma_k)\theta^m)) - T^* + \theta^m UL^*\right)\right)\right] +$$

$$+V(\hat{Z}) - V(Z^*) \leq 0$$

By definition of $\hat{Z}$ it must be the case that $V(\hat{Z}) - V(Z^*) \geq 0$. In addition notice that $\theta^m > (1-\beta)\left(\frac{\alpha}{C_0^*}(R_0^*(\gamma_k+(K_{-1}-\gamma_k)\theta^m)) - T^* + \theta^m UL^*\right)$. If not, we would have

$$\theta^m \leq (1-\beta)(\frac{\alpha}{C_0^*}(R_0^*(\gamma_k+(K_{-1}-\gamma_k)\theta^m)) - T^* + \theta^m UL^*)$$

or

$$\theta^m\left[1-(1-\beta)UL^* - (1-\beta)\frac{\alpha}{C_0^*}R_0^*(K_{-1}-\gamma_k)\right] \leq (1-\beta)\left[-T^* + \frac{\alpha}{C_0^*}\left(R_0^*\gamma_k\right)\right]$$

Condition (4) in Lemma 2 implies that $(1-\beta)(\frac{\alpha}{C_0^*}(R_0^*(\gamma_k+(K_{-1}-\gamma_k)) - T^* + UL^*) = 1$. Replacing this condition in the inequality above yields:

$$\theta^m(1-\beta)(-T^*) \leq (1-\beta)(-T^*)$$

But since $T^* > 0$ the above inequality implies $\theta^m \geq 1$, a contradiction. Therefore, $\theta^m > (1-\beta)\left(\frac{\alpha}{C_0^*}(R_0^*(\gamma_k+(K_{-1}-\gamma_k)\theta^m) - T^*) + \theta^m UL^*\right)$. This last strict inequality implies $P(\widehat{Z}) - P(M) > 0$, a contradiction ■

17.4 Proof of Lemma 4:

Proof. The following slightly modifies the proof in Bassetto and Benhabib (2006). If the claim is not true, then $\{C^*(s^t), K^*(s^t)\}_{s^t > s^{\tilde{t}}}$ does not satisfy the first order conditions in the following problem:

$$\max_{\{C(s^t),K(s^t)\}_{s^t>s^{\tilde{t}}}} \sum_{s^t>s^{\tilde{t}}} \beta^t \Pr(s^t)\left[\alpha \log(C(s^t)) + (1-\alpha)\log(1-L(s^t))\right]$$

$$\text{s.t.} = \begin{cases} C(s^t) + K(s^t) + g(s^t) \leq F\Big(L(s^t), K(s^{t-1}), s^t\Big) + (1-\delta)K(s^{t-1}) \text{ for all } s^t > s^{\tilde{t}} \\ \frac{1}{1-\beta} = \mathrm{E}(\widetilde{W}_0(\theta, T, \tau_{k0}))\frac{\alpha}{C_0} + UL \\ K^*(s^{\tilde{t}}),\ K_{-1},\ T^*,\ \tau_{k0} \text{ and } \{L^*(s^t)\}_{s^t>s^{\tilde{t}}} \text{ given} \end{cases}$$

Then it follows there must be an alternative allocation $\{\hat{C}(s^t), \hat{K}(s^t)\}_{s^t>s^{\tilde{t}}}$ satisfying the constraints above that yields a higher value for the return function.

Let $s^t \not> s^{\tilde{t}}$ with $t > \tilde{t}$ denote a history s^t that does **not** follow the history $s^{\tilde{t}}$.

Since the utility for the median type is increasing in the value of the utility for the mean type, it follows that $\{\hat{C}(s^t), \hat{K}(s^t)\}_{s^t>s^{\tilde{t}}}$ and K_{-1}, T^*, τ_{k0}, $\{L^*(s^t)\}_{t\geq 0}$, $\{C^*(s^t), K^*(s^t)\}_{t<\tilde{t}}$ and $\{C^*(s^t), K^*(s^t)\}_{s^t \not> s^{\tilde{t}}}$ is a feasible allocation for the median voter's problem that improves the objective function, a contradiction □

17.5 Proof of Lemma 5:

Proof. By Lemma 4 constraint (NN) can be disregarded. Let μ be the multiplier associated with the constraint on τ_{k0}. Then the Foc for τ_{k0} generates:

$$\frac{(1-\theta^m)}{\varphi(\theta^m)} = \mu \frac{C_0}{\alpha(K_{-1} - \gamma_k)} \frac{1}{(F_{k_0} - \delta)} \tag{19}$$

Therefore $\mu > 0$, which implies that there is a corner solution for τ_{k0}. Next, define $R_0 = 1 + (1 - \bar{\tau})(F_{k_0} - \delta)$. The first order conditions without considering the conditions in (UB) imply:

$$\frac{1}{C_0^2} R_0 (K_{-1} - \gamma_k) \frac{(1-\theta^m)}{\varphi(\theta^m)} + \frac{1}{C_0} = \beta E \left[\frac{1 - \delta + F_k(s^1)}{C(s^1)} \mid s_0 \right]$$

Constraint (UB) at s_0 is satisfied when:

$$\frac{1}{C_0^2} R_0 (K_{-1} - \gamma_k) \frac{(1-\theta^m)}{\varphi(\theta^m)} - E \left[\frac{\bar{\tau}(F_k(s^1) - \delta)}{C(s^1)} \mid s_0 \right] \leq 0 \tag{20}$$

Because of the log utility function on consumption, any solution will have $C_0^* > 0$ regardless the size of $K_{-1} - \gamma_k \geq 0$. From Lemma 3 we have that $\theta^m < \varphi(Z^*, \theta^m) < 1$. Therefore the above holds when $\delta = 0$ and $K_{-1} - \gamma_k$ is arbitrarily close to zero.As it is well known from Chari and Kehoe (1999), the process for taxes on capital income as a function of implementable allocations is not uniquely determined. In particular, as in Werning (2006), one such process can be constructed by:

$$\frac{1 + (1 - \tau_k(s^1))(F_k(s^1) - \delta)}{1 - \delta + F_k(s^1)} = \frac{U_c(s_0)}{V_{c0}(\theta^m, Z)} \frac{V_{c1}(\theta^m, Z)}{U_c(s^1)} \text{ for all } s^1$$

where $V(\cdot)$ stands for the objective function in the median voter's problem. Then using (19) the expression yields:

$$\frac{1 + (1 - \tau_k(s^1))(F_k(s^1) - \delta)}{1 - \delta + F_k(s^1)} = \frac{1/C_0}{\left[\mu \frac{R_0}{\alpha C_0 (F_{k_0} - \delta)}\right] + 1/C_0} \text{ for all } s^1$$

Clearly, the above equation implies that $1 + (1 - \tau_k(s^1))(F_k(s^1) - \delta) < 1 - \delta + F_k(s^1)$, and in turn that $\tau_k(s^1) > 0$ for all s^1. This proves the second line.

Finally, the Foc's without considering UB satisfy the constraint for all $t \geq 2$. Furthermore, the implied taxes on capital returns are zero $\square$

17.6 Median Voter Result and Characterization in the General Case

As in the main part of the paper, set $\varphi(\theta) \equiv 1/\lambda(\theta)$, where $\lambda(\theta)$ is the multiplier related to the present value budget constraint of type θ.

Working with the first order conditions with respect to individual consumption and labor yields:

$$c(s^t,\theta) = \frac{\theta^{\frac{-(1-\alpha)(1-\sigma)}{\sigma}}\varphi^{1/\sigma}(\theta)}{\int_\Theta \theta^{\frac{-(1-\alpha)(1-\sigma)}{\sigma}}\varphi^{1/\sigma}(\theta)f(\theta)d\theta}C(s^t) = \omega_c(\theta)C(s^t)$$

$$1 - \frac{l(s^t,\theta)}{\theta} = \frac{\theta^{\frac{\alpha(1-\sigma)-1}{\sigma}}\varphi^{1/\sigma}(\theta)}{\int_\Theta \theta^{\frac{-(1-\alpha)(1-\sigma)}{\sigma}}\varphi^{1/\sigma}(\theta)f(\theta)d\theta}[1-L(s^t)] = \omega_L(\theta)[1-L(s^t)]$$

and

$$\begin{aligned} p(s^t) &= \alpha\Phi^\sigma[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}]^{-\sigma}\left[\frac{C(s^t)}{1-L(s^t)}\right]^{\alpha-1} \\ p(s^t)w(s^t)(1-\tau_l(s^t)) &= (1-\alpha)\Phi^\sigma[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}]^{-\sigma}\left[\frac{C(s^t)}{1-L(s^t)}\right]^{\alpha} \end{aligned}$$

where $\Phi = \int_\Theta \theta^{\frac{-(1-\alpha)(1-\sigma)}{\sigma}}\varphi^{1/\sigma}(\theta)f(\theta)d\theta$.

Let $U^{CE}(C(s^t),L(s^t);\varphi) = \frac{\left[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}\right]^{1-\sigma}}{1-\sigma}$. As in the logarithmic case, replacing prices and individual allocations in the budget constraint for each agent θ yields:

$$\sum_t \beta^t \Pr(s^t)\omega_c(\theta)\alpha(1-\sigma)U^{CE}(C(s^t),L(s^t)) = U^{CE}_{c0}\widetilde{W}_0(\theta,T,\tau_{k0})$$
$$+\sum_t \beta^t \Pr(s^t)(1-\alpha)\left[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}\right]^{-\sigma}\left(\frac{C(s^t)}{(1-L(s^t))}\right)^\alpha \theta\left[1-\omega_L(\theta)\left(1-L(s^t)\right)\right]$$

Let $UL = \sum_t \beta^t \Pr(s^t) u\Big(C(s^t), 1 - L(s^t)\Big)\frac{(1-\alpha)(1-\sigma)}{1-L(s^t)}$, then we can write

$$(1-\sigma)\omega_c(\theta)\sum_t \beta^t Pr(s^t)U^{CE}(C(s^t), L(s^t)) = U_{c0}^{CE}\widetilde{W}_0(\theta, T, \tau_{k0}) + \theta UL$$

Where we have used the fact that $\omega_c(\theta) = \theta\omega_L(\theta)$, then

$$(1-\sigma)\frac{\varphi(\theta)^{1/\sigma}}{\Phi}\theta^{\frac{-(1-\alpha)(1-\sigma)}{\sigma}}V(Z) = U_{c0}^{CE}\widetilde{W}_0(\theta, T, \tau_{k0}) + \theta UL$$

where $V(Z) = \sum_t \beta^t Pr(s^t)U^{CE}(C(s^t), L(s^t))$.

Therefore

$$\frac{\varphi(\theta)^{1/\sigma}}{\Phi\theta^{\frac{-(1-\alpha)(1-\sigma)}{\sigma}}} = \frac{[U_{c0}^{CE}\widetilde{W}_0(\theta, T, \tau_{k0}) + \theta UL]}{(1-\sigma)V(Z)} \geq 0 \tag{21}$$

The utility of any agent in a particular competitive equilibrium is given by:

$$V(Z;\theta) = \frac{\sum_t \beta^t Pr(s^t)\left\{[\omega_c(\theta)C(s^t)]^\alpha[\omega_L(\theta)\left(1 - L(s^t)\right)]^{1-\alpha}\right\}^{1-\sigma}}{1-\sigma} = \frac{\left(\omega_c(\theta)^\alpha\omega_L(\theta)^{1-\alpha}\right)^{1-\sigma}}{\Phi^\sigma}V(Z)$$

but notice that $\left(\omega_c(\theta)^\alpha\omega_L(\theta)^{1-\alpha}\right)^{1-\sigma} = \frac{\varphi(\theta)^{\frac{1-\sigma}{\sigma}}}{\Phi^{1-\sigma}}\theta^{-\frac{(1-\alpha)(1-\sigma)}{\sigma}}$. So we have:

$$V(Z;\theta) = \frac{\varphi(\theta)^{\frac{1-\sigma}{\sigma}}}{\Phi}\theta^{-\frac{(1-\alpha)(1-\sigma)}{\sigma}}V(Z) \tag{22}$$

Finally, substituting (21) into (22) generates

$$V(Z;\theta) = \frac{[U_{c0}^{CE}\widetilde{W}_0(\theta, T, \tau_{k0}) + \theta UL]^{1-\sigma}}{\theta^{(1-\alpha)(1-\sigma)}[(1-\sigma)V(Z)]^{1-\sigma}}V(Z) \tag{23}$$

Given two competitive equilibrium allocations $Z, \hat{Z} \in \Upsilon$, type θ prefers Z to $\hat{Z}$ iff $V(Z;\theta) \geq V(\hat{Z};\theta)$, or alternatively, $\log(V(Z;\theta)/V(\hat{Z};\theta)) \geq 0$.

Equation (23) can be used to compute the ratio $V(Z;\theta)/V(\hat{Z},\theta)$ as

$$\frac{V(Z;\theta)}{V(\hat{Z};\theta)} = \left(\frac{[U_{c0}^{CE} W_0(\theta, T, \tau_{k0}) + \theta UL]}{[\widetilde{U}_{c0}^{CE} W_0(\theta, \hat{T}, \hat{\tau}_{k0}) + \theta \hat{U}L]}\right)^{1-\sigma} \left(\frac{\hat{\Phi}V(Z;\theta)}{\Phi V(\hat{Z};\theta)}\right)^{\sigma}$$

Proposition 1.1: **(MVT - Inequality in both labor skills and initial wealth)**

Suppose the initial wealth is an affine function of skills, i.e., $k_0(\theta) = \nu_1 + \nu_2\theta$. Consider any $Z, \hat{Z} \in \Xi$. If $\theta^m \in S_{Z,\hat{Z}}$, then either $[\underline{\theta}, \theta^m] \subseteq S_{Z,\hat{Z}}$ or $[\theta^m, \overline{\theta}] \subseteq S_{Z,\hat{Z}}$.

Proof. $W_0(\hat{T}, \hat{\tau}_{k0})$ can be written as $W_0(\hat{T}, \hat{\tau}_{k0}) = a + R\nu_2\theta$. Then consider the following derivative

$$\begin{aligned}
\frac{\partial \log\left(\frac{V(Z;\theta)}{V(\hat{Z};\theta)}\right)}{\partial \theta} &= (1-\sigma)\left[\frac{UL + \nu_2 R U_{c0}^{CE}}{[U_{c0}^{CE} W_0(T, \tau_{k0}) + \theta UL]} - \frac{\hat{U}L + \hat{R}\nu_2 \hat{U}_{c0}^{CE}}{[\hat{U}_{c0}^{CE} W_0(\hat{T}, \hat{\tau}_{k0}) + \theta \hat{U}L]}\right] \\
&= (1-\sigma)\left[\frac{\hat{a}\hat{U}_{c0}^{CE}(UL + R\nu_2 U_{c0}^{CE}) - aU_{c0}^{CE}(\hat{U}L + \hat{R}\nu_2 \hat{U}_{c0}^{CE})}{[U_{c0}^{CE} W_0(T, \tau_{k0}) + \theta UL][\hat{U}_{c0}^{CE} W_0(\hat{T}, \hat{\tau}_{k0}) + \theta \hat{U}L]}\right]
\end{aligned}$$

Therefore, as in the log case, the sign of the derivative does not depend on θ. □

17.6.1 Characterization

The objective function for the median voter problem is given by:

$$V(Z, \theta^m) = V(Z)\left\{\frac{u_{c0}[\widetilde{W}_0(\theta^m, T, \tau_{k0}) - E[\widetilde{W}_0(\theta, T, \tau_{k0})]] + (\theta^m - 1)UL + (1-\sigma)V(Z)}{[(1-\sigma)V(Z)]}\right\}^{1-\sigma} \tag{24}$$

where $V(Z) = \sum_t \beta^t Pr(s^t) U^{CE}(C(s^t), L(s^t))$ and $UL = \sum_t \beta^t \Pr(s^t) u\Big(C(s^t), 1 - L(s^t)\Big) \frac{(1-\alpha)(1-\sigma)}{1-L(s^t)}$.

In the general case, problem P(M) becomes:

$$\max_{\{C,L,K,T,\tau_0\}} V(Z,\theta^m)$$

$$\text{s.t.} \begin{cases} C(s^t)+K(s^t)+g(s^t) \leq F\Big(L(s^t),K(s^{t-1}),s^t\Big)+(1-\delta)K(s^{t-1}) \quad (\mathbf{RC}); \\ U_c(s^t) \geq \beta \sum_{s_{t+1}} \Pr(s_{t+1}|s^t)U_c(s^{t+1})[1+(1-\overline{\tau})(F_k(s^{t+1})-\delta)] \qquad (\mathbf{UB}) \\ U_{c0}^{CE}[\widetilde{W}_0(\underline{\theta},T,\tau_{k0})-E[\widetilde{W}_0(\theta,T,\tau_{k0})]]+(\underline{\theta}-1)UL+(1-\sigma)V(Z) \geq 0 \quad (\mathbf{NN}) \\ (1-\sigma)V(Z)=U_{c0}^{CE}E[\widetilde{W}_0(\theta,T,\tau_{k0})]+UL \\ \tau_{k0} \leq \overline{\tau} \end{cases}$$

The return function in the problem above can be written as $V(Z,\theta^m)=[\chi(Z,\theta^m)]^{1-\sigma}V(Z)$, where $\chi(Z,\theta^m)^{1-\sigma}$ is the first part of (24):

$$\chi(Z,\theta^m) \equiv \frac{U_{c0}[\widetilde{W}_0(\theta^m,T,\tau_{k0})-E[\widetilde{W}_0(\theta,T,\tau_{k0})]]+(\theta^m-1)UL+(1-\sigma)V(Z)}{(1-\sigma)V(Z)}$$

It can be shown that the partial derivatives are given by:

$$\hat{V}_C(s^t)=\beta^t\Pr(s^t)[\chi(Z,\theta^m)]^{1-\sigma}\left\{\frac{(1-\sigma)}{\chi(Z,\theta^m)}\left[\frac{(\theta^m-1)(1-\alpha)}{1-L(s^t)}+1\right]+\sigma\right\}U_c(s^t) \tag{25}$$

$$\hat{V}_L(s^t)=\beta^t\Pr(s^t)[\chi(Z,\theta^m)]^{1-\sigma}\left\{\frac{(\theta^m-1)[(1-\alpha)(1-\sigma)-1]}{[1-L(s^t)]\chi(Z,\theta^m)}+\frac{1-\sigma}{\chi(Z,\theta^m)}+\sigma\right\}U_L(s^t) \tag{26}$$

With some abuse of notation, let:

$$a(s^t)=[\chi(Z,\theta^m)]^{1-\sigma}\left\{\frac{(1-\sigma)}{\chi(Z,\theta^m)}\left[\frac{(\theta^m-1)(1-\alpha)}{1-L(s^t)}+1\right]+\sigma\right\}, \quad b(\theta^m)=\frac{-(\theta^m-1)}{(1-L(s^t))\chi(Z,\theta^m)}$$

Lemma 9: If $\theta^m<1$ then in any solution to P(M) we have $\theta^m \leq \chi(Z,\theta^m)<1$.

Proof. Clearly $\chi(Z,\theta^m)<1$ when $\theta^m<1$, so we only need to show that $\theta^m<\chi(Z,\theta^m)$. First, without lost of generality assume that $b_0(\theta)=0\forall\theta$. Then $W_0(\theta,T,\tau_{k0})=$

$\gamma_k + (K_0 - \gamma_k)\theta - T$ because $R_0 = 1$ at the optimum. As in the proof of Lemma 4, in the solution to median voter problem (with $\theta^m < 1$) we have $T \leq 0$. Then it must be true that:

$$
\begin{aligned}
0 &\leq U_{c0}\left[-(\theta^m - 1)\gamma_k + (\theta^m - 1)T\right] \\
&= U_{c0}\left[(\theta^m - 1)(K_0 - \gamma_k) - (\theta^m - 1)K_0 + (\theta^m - 1)T\right] \\
&= U_{c0}\left[W_0(\theta^m, T, \tau_{k0}) - E[W_0(\theta, T, \tau_{k0})]\right] - (\theta^m - 1)U_{c0}E[W_0(\theta, T, \tau_{k0})] \\
&= U_{c0}\left[W_0(\theta^m, T, \tau_{k0}) - E[W_0(\theta, T, \tau_{k0})]\right] - (\theta^m - 1)\big[U_{c0}E[W_0(\theta, T, \tau_{k0})] + UL\big] + (\theta^m - 1)UL \\
&= U_{c0}\left[W_0(\theta^m, T, \tau_{k0}) - E[W_0(\theta, T, \tau_{k0})]\right] - (\theta^m - 1)(1 - \sigma)V + (\theta^m - 1)UL
\end{aligned}
$$

Where $(1 - \sigma)V = U_{c0}E[W_0(\theta, T, \tau_{k0})] + UL$ from the **MKT** constraint. The last inequality can be written as

$$
U_{c0}\left[W_0(\theta^m, T, \tau_{k0}) - E[W_0(\theta, T, \tau_{k0})]\right] + (\theta^m - 1)UL \geq (\theta^m - 1)(1 - \sigma)V
$$

or,

$$
\frac{U_{c0}\left[W_0(\theta^m, T, \tau_{k0}) - E[W_0(\theta, T, \tau_{k0})]\right] + (\theta^m - 1)UL}{(1 - \sigma)V} + 1 \geq \theta^m
$$

But the left hand side of the last inequality is simply $\chi(Z, \theta^m)$. $\square$

Lemma 10: If $\theta^m < 1$ and $1 < \sigma \leq \frac{1}{1-\theta^m}$, then in any solution to P(M) we have $a(s^t) > 0$ for all s^t.

Proof. First, consider the case $\sigma > 1$. $a(s^t)$ is greater than zero as long as:

$$\frac{(1-\sigma)}{\chi(Z,\theta^m)}\left[\frac{(\theta^m-1)(1-\alpha)}{1-L(s^t)}+1\right]+\sigma > 0$$

or

$$\frac{(\theta^m-1)(1-\alpha)}{1-L(s^t)}+1 < \frac{\sigma\chi(Z,\theta^m)}{\sigma-1}$$

Since $(\theta^m - 1) < 0$ the inequality above is indeed true as long as $\sigma \leq \frac{1}{1-\chi(Z,\theta^m)}$. But since by Lemma 9 $\frac{1}{1-\theta^m} < \frac{1}{1-\chi(Z,\theta^m)}$, $\sigma \leq \frac{1}{1-\theta^m}$ is a sufficient condition □

Lemma 11(The Bang-Bang Property:) Assume $1 < \sigma \leq \frac{1}{1-\theta^m}$. In the solution for the median voter's problem, if there exists $\tilde{t}$ such that the implied tax $\tau_k(s^{\tilde{t}}) < \overline{\tau}$ for all $s^{\tilde{t}}$ then

$$U_C(s^t) = \beta\sum_{s_{t+1}}\Pr(s_{t+1}|s^t)U_C(s^{t+1})[1+F_k^*(s^{t+1})-\delta] \;\; \forall \;\; t \geq \tilde{t}$$

and therefore $\tau_k(s^{\tilde{t}}) = 0$ for all $t \geq \tilde{t}$.

We omit the proof here, since it uses the same reasoning as the proof of Lemma 5. The key element of the proof is that the return function in the median voter's problem is increasing in both aggregate consumption and leisure when $1 < \sigma \leq \frac{1}{1-\theta^m}$.

Lemma 12:(The Capital Tax Result) Suppose $1 < \sigma \leq \frac{1}{1-\theta^m}$, $\theta^m < E(\theta) = 1$, $k_0(\theta) = \gamma_k + (K_0 - \gamma_k)\theta$ with $K_0 - \gamma_k > 0$, and $b_0(\theta) = \hat{b_0} \; \forall \; \theta \; \in \; \Theta$. Then there exists $\hat{t} > 1$ such that

$$\tau_k(s^t) = \begin{cases} \overline{\tau} = 1 \text{ for } t < \hat{t} \\ 0 \leq \tau_k(s^t) < \overline{\tau} \text{ for } t = \hat{t} \\ 0 \text{ for all } s^t \text{ such that } t > \hat{t} \end{cases}$$

The proof is standard, dating from the original work of Chamley (1986). The condition $1 < \sigma \leq \frac{1}{1-\theta^m}$ ensures that the median voter's value function is increasing in aggregate consumption, and therefore it cannot be the case that the constraint UB is always binding when there is discounting. Otherwise the standard reasoning would not apply.

Proposition 4. ***(Labor Tax Result)*** *Suppose $\sigma \leq \frac{1}{1-\theta^m}$ and $\theta^m < E(\theta)$. Then there exists $\hat{t} > 1$ such that, for $t \geq \hat{t}$:*

1. $0 < \tau_l(s^t) < 1$.
2. *$\tau_l(s^t)$ depends on s^t only through $L(s^t)$.*
3. *$\tau_l(s^t)$ is strictly increasing in $[1 - \theta^m]$.*

Proof. Case 1: $1 < \sigma \leq \frac{1}{1-\theta^m}$

Because of Lemma 12, and since (NN) is not binding, the first order condition with respect to labor is (for $t \geq \hat{t}$):

$$-\frac{V_{Lt}(Z;\theta)}{V_{ct}(Z;\theta)} = F_L(s^t) \tag{27}$$

¿From the CE we know that

$$1-\tau_l(s^t) = -\frac{U_L(s^t)}{F_L(s^t)U_c(s^t)}$$

Combining the last two equations and using (25) and (26) generates

$$1-\tau_l(s^t) = \frac{a(s^t)}{a(s^t)+b(\theta^m)}$$

Thus, if $1 < \sigma \leq \frac{1}{1-\theta^m}$, by Lemma 10 we have $a(s^t) > 0$, and therefore $0 < 1-\tau_l(s^t) < 1$ for all s^t.

Case 2: $0 < \sigma < 1$.

Suppose that the constraint UB is not binding for all $t \geq \hat{t}$. Later we will check that constraints. We can write $\tau_l(s^t)$ as:

$$\tau_l(s^t) = \frac{-(\theta^m - 1)}{(\theta^m - 1)[(1-\alpha)(1-\sigma)-1] + [1-L(s^t)][(1-\sigma)+\sigma\chi(Z,\theta^m)]} \qquad (28)$$

Which implies that $\tau_l(s^t) > 0$ when $0 < \sigma < 1$. In this case, $\tau_l(s^t) < 1$ follows from the intratemporal first order condition in the competitive equilibrium. Otherwise, the marginal productivity of labor should be negative.

Next we claim that, if $\theta^m < 1$ and $0 < \sigma < 1$ then $a(s^t) > 0$ for all s^t.

First, notice that $\tau_l(s^t) < 1$ implies that $a(s^t)$ and $a(s^t)+b(\theta^m)$ must have the same sign. $b(\theta^m) > 0$ with $\tau_l(s^t) > 0$ implies the claim. Finally, since $a(s^t) > 0$ for all $t \geq \hat{t}$, constraint UB is not binding for high enough t $\square$

17.7 Numerical Algorithm

We numerically approximate the solution to the following problem:

$$\max_{\omega_c(\theta^m),C,L,K} \left(\frac{\omega_c(\theta^m)}{\theta^{m1-\alpha}}\right)^{1-\sigma} \sum_{t,s^t} \beta^t \Pr(s^t) \frac{\left[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}\right]^{1-\sigma}}{1-\sigma}$$

$$\text{s.t.} \begin{cases} \omega_c(\theta^m) \leq 1 + \frac{(\theta^m-1)\sum_{t,s^t}\rho(s^t)U_L(s^t)}{(1-\sigma)V(Z)} \\ \text{Resource constraint} \\ \text{non-negativity constraints} \\ L(s^t) \leq 1 \end{cases}$$

where

$$U_L(s^t) \equiv \beta^t \Pr(s^t)\Big(C(s^t)^\alpha(1-L(s^t))^{1-\alpha}\Big)^{1-\sigma} \frac{(1-\alpha)}{1-L(s^t)}$$

and

$$V(Z) \equiv \textstyle\sum_{t,s^t} \beta^t \Pr(s^t) \frac{\left[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}\right]^{1-\sigma}}{1-\sigma}.$$

A straightforward extension of Lemma 9 in order to allow for stochastic labor skills shows that, in the solution to the problem above, $\omega_c(\theta^m) \in [\theta^m, 1)$. Let λ be the multiplier related to the first constraint above, which clearly binds in the solution. Let $\xi(s^t)$ be the multiplier on the resource constraint at history s^t. Then the Lagrangean is given by:

$$\mathcal{L} = \sum_{t,s^t} \Pr(s^t)\beta^t \left[\left(\frac{\omega_c(\theta^m)}{\theta^{m1-\alpha}}\right)^{1-\sigma} U(C(s^t), L(s^t))\right] + \lambda\left[(\theta^m-1)\sum_{t,s^t}\rho(s^t)U_L(s^t) + (1-\sigma)V(Z)(1-\omega_c(\theta^m))\right] +$$
$$\sum_{t,s^t} \xi(s^t)\left[F\Big(L(s^t), K(s^{t-1}), s^t\Big) + (1-\delta)K(s^{t-1}) - C(s^t) - K(s^t) - g(s^t)\right]$$

where $U(C(s^t), L(s^t)) \equiv \frac{\left[C(s^t)^\alpha(1-L(s^t))^{1-\alpha}\right]^{1-\sigma}}{1-\sigma}$.

We then can rewrite $\mathcal{L}$ as:

$$\mathcal{L} = \sum_{t,s^t} \Pr(s^t)\beta^t \Bigg[\left(\frac{\omega_c(\theta^m)}{\theta^{m\,1-\alpha}}\right)^{1-\sigma} U(C(s^t), L(s^t)) + \\ \lambda(1-\sigma)U(C(s^t), L(s^t)) \left(1 - \omega_c(\theta^m) + \rho(s^t)(\theta^m - 1)\frac{1-\alpha}{1-L(s^t)}\right)\Bigg] + \\ \sum_{t,s^t} \xi(s^t) \left[F\left(L(s^t), K(s^{t-1}), s^t\right) + (1-\delta)K(s^{t-1}) - C(s^t) - K(s^t) - g(s^t)\right]$$

Taking λ and $\omega_c(\theta^m)$ as given, there exists a functional equation problem (FEP) with a modified return function that solves $\mathcal{L}$ above. Such return function is given by:

$$\widehat{U}(C(s^t), L(s^t); \lambda, \omega_c(\theta^m)) \equiv \\ \left(\frac{\omega_c(\theta^m)}{\theta^{m\,1-\alpha}}\right)^{1-\sigma} U(C(s^t), L(s^t)) + \lambda(1-\sigma)U(C(s^t), L(s^t)) \left(1 - \omega_c(\theta^m) + \rho(s^t)(\theta^m - 1)\frac{1-\alpha}{1-L(s^t)}\right)$$

Denote $V(K; \lambda, \omega_c(\theta^m))$ the unique function solving the FEP. Using the product topology in the problem in question, we can apply Theorem 3 in Milgrom and Segal (2002). By setting $\frac{\partial V(K;\lambda,\omega)}{\partial \omega_c(\theta^m)} = 0$ we get

$$\omega_c(\theta^m) = [\theta^{(1-\alpha)(1-\sigma)}\lambda]^{-1/\sigma}$$

The numerical solution then uses a two step algorithm. First, for a given λ, and therefore $\omega_c(\theta^m)$ from the equation above, we solve the FEP using value function iteration for a grid of 300 points for the capital stock. In the second step, for each capital stock, we do a grid with 100 points for λ and find $\lambda^*(K)$ that attains

$\frac{\partial V(K;\lambda,\omega)}{\partial \lambda} = 0$. Because λ and $\omega_c(\theta^m)$ are related by an equation and $\omega_c(\theta^m) \in [\theta^m, 1)$, we can reduce the size of the grid for $\lambda' s$ in a great extent.

We check the numerical solution by evaluating the analytic first-order conditions from the original problem.

Part V

Appendix B: proofs of Chapter III

17.8 Proof of Lemma 13

(Only If) Let $\widetilde{a}^i_{t+1} \equiv a^i_{t+1} - \underline{a}$. The necessary conditions for the maximization of the individual's problem are :

$$u^i_{ct} = \lambda^i_t,\ u^i_{lt} = \lambda^i_t p_{lt},\ (\lambda^i_t - \lambda^i_{t+1}(1 - \tau_{k_{t+1}})r_{t+1})\widetilde{a}^i_{t+1}0$$

and the following TVC condition:

$$\lim_{t\to\infty} a^i_{t+1}\lambda^i_t = 0$$

Using the both the TVC and the Foc's we get the present value budget (3) for each i. The other conditions are trivial. Notice that the Foc's can be manipulated such that individual allocations of labor and consumption can be expressed as shares of the aggregates. Moreover, $(1 - \tau_{lt}) = \frac{1-\alpha}{\alpha}\frac{C_t}{1-L_t}\frac{1}{F^{\xi}_l}$.

(If) Set $(1 - \tau_{lt}) = \frac{1-\alpha}{\alpha}\frac{C_t}{1-L_t}\frac{1}{F^{\xi}_l}$, $r_t = F_{kt}$. Rental prices for labor and bond prices are given by $w_t = F_{lt}$.

Taxes on capital are given by $1 - \frac{u^i_{ct}}{u^i_{c_{t+1}} r_{t+1}}$. By setting $\lambda^i_t = u^i_{ct}$ the foc's are met. The TVC for bonds holds because the bounds on borrowing and lending.

By construction the foc's are met. Construct total wealth levels such that the budget constraint is satisfied with equality in each period:

$$\lambda_t^i(r_t(1-\tau_{kt})a_t^i) = \sum_{s=t}^{\infty} \lambda_s^i[(1-\tau_{ls})w_s l_s^i + T_s - c_s^i]$$

The resource constraint is met by integrating the budget(which holds with equality) over types in each period.

17.9 Proof of Lemma 14

Suppose (1)-(4) characterize a CE. Then set $\varphi^i = [\lambda_t^i]^{-1}[\beta^t \frac{\alpha}{C_t}]$ and $[c_t^i, 1-l_t^i] = \varphi^i[C_t, 1-L_t]$. Then using these expressions, we get (2). Moreover, using these expressions in (3) we get the expression φ^i in (3'). It remains to show that $\int_I \varphi^i dF(i) = 1$. The construction of prices and taxes from the foc's(which are satisfied using φ^i) gives:

$$(1-L_t)(1-\tau_{lt})F_{lt} = \frac{1-\alpha}{\alpha}C_t$$

Then:

$$\begin{aligned}
c_t &= \alpha[(C_t - F_{lt}L_t) + \tau_{lt}F_{lt}L_t + (1-\tau_{lt})F_{lt}] \\
c_t &= \alpha[(-K_{t+1} + F_{kt}K_t) \pm \tau_{kt}F_{kt}K_t + \tau_{lt}F_{lt}L_t + (1-\tau_{lt})F_{lt}] \\
c_t &= \alpha[-K_{t+1} + (1-\tau_{kt})F_{kt}K_t + T_t + (1-\tau_{lt})F_{lt}] \\
\beta^t &= \beta^t \frac{\alpha}{C_t}[-K_{t+1} + (1-\tau_{kt})F_{kt}K_t + T_t + (1-\tau_{lt})F_{lt}]
\end{aligned}$$

Notice that we have used the fact that $T_t = \tau_{kt}F_{kt}K_t + \tau_{lt}F_{lt}L_t$. Since the final expression above is true in each period, we have (using $\left[\frac{\alpha\beta^t}{C_t} - \frac{\alpha\beta^{t+1}}{C_{t+1}}(1-\tau_{k_{t+1}})F_{k_{t+1}})\right] K_{t+1} =$ 0):

$$\frac{1}{1-\beta} = \frac{\alpha}{C_0}(1-\tau_{k0})F_{k0}K_0 + \sum_{t\geq 0}\beta^t \frac{\alpha}{C_t}[(1-\tau_{lt})F_{lt} + T_t]$$

or

$$1 = (1-\beta)\left[\frac{\alpha}{C_0}(1-\tau_{k0})F_{k0}K_0 + \sum_{t\geq 0}\beta^t \frac{\alpha}{C_t}[(1-\tau_{lt})F_{lt} + T_t]\right]$$

But this is equivalent to $\int_I \varphi^i dF(i) = 1$.

The converse is clearly true by setting $\lambda_t^i = \frac{\alpha\beta^t}{\varphi^i C_t}$ □

17.10 Proof of Lemma 16

By induction. Here we show only that $\varphi(\{Z_t^*, \tau_t\}_{t\geq 0}, k_0^i) = \varphi(\{Z_t^*, \tau_t\}_{t\geq 1}, a_1^{i*})$.

Let $(\varphi^i \equiv \varphi(\{Z_t^*, \tau_t\}_{t\geq 0}, k_0^i))$. Using the foc's in a CE at time zero (in particular: $\frac{\alpha}{C_0} = \beta\frac{\beta\alpha}{C_1}(1-\tau_{k1})r_1$):

$$\begin{aligned}
\frac{\alpha\beta}{C_1}(1-\tau_{_{k1}})(a_1^i) &= \frac{\alpha}{C_0}[(1-\tau_{_{k0}})F_{k0}k_0^i + (1-\tau_{l0})F_{l0}(1-\varphi^i(1-L_0)) + T_0 - \varphi^i C_0] \\
&= \frac{\alpha}{C_0}[(1-\tau_{_{k0}})F_{k0}k_0^i + T_0 + (1-\tau_{l0})F_{l0}] - \alpha\varphi^i - \frac{\alpha(1-\tau_{l0})F_{l0}\varphi^i(1-L_0)}{C_0} \\
&= \frac{\alpha}{C_0}[(1-\tau_{_{k0}})F_{k0}k_0^i + T_0 + (1-\tau_{l0})F_{l0}] - \alpha\varphi^i - (1-\alpha)\varphi^i
\end{aligned}$$

where in the last line we use the intratemporal foc.

Then we can write $\varphi(\{Z_t^*, \tau_t\}_{t\geq 1}, a_1^{i*})$:

$$
\begin{aligned}
\beta\varphi(\{Z_t^*, \tau_t\}_{t\geq 1}, a_1^{i*}) &= (1-\beta)\left[\beta\frac{\alpha}{C_1}(1-\tau_{k1})(F_{k1}a_1^i) + \sum_{t\geq 1}\frac{\alpha\beta^t}{C_n}((1-\tau_{ln})F_{ln} + T_n)\right] \\
&= (1-\beta)\left[\frac{\alpha}{C_0}(1-\tau_{k0})F_{k0}k_0^i + \sum_{t\geq 0}\frac{\alpha\beta^t}{C_n}((1-\tau_{ln})F_{ln} + T_n) - \varphi(\theta^i, \{Z_t^*, \tau_t\}_{t\geq 0}, k_0^i))\right] \\
&= \varphi(\{Z_t^*, \tau_t\}_{t\geq 0}, k_0^i)) - (1-\beta)\varphi(\{Z_t^*, \tau_t\}_{t\geq 0}, k_0^i)) \\
&= \beta\varphi(\{Z_t^*, \tau_t\}_{t\geq 0}, k_0^i)) \blacksquare
\end{aligned}
$$

17.11 Proof of Lemma 19

We first assume that, in the solution, the implied taxes on capital are less then the upper bound at t=1. Then by a known result all future taxes on capital are zero and therefore all multipliers associated with the second constraint are zero. Later we will use the condition in the Lemma to show that indeed $\tau_{k1} < \overline{\tau}_k$.

The Foc's assuming $\psi_t = 0 \ \forall \ t$ yield:

$$\frac{\alpha}{C_0}\left[1 - \frac{\gamma}{C_0}\frac{(1-\overline{\tau}_k)F(K_0, L_0)(\theta^m - 1)}{\varphi^m}\right] = \lambda_0 = \frac{\alpha\beta}{C_1}F_{k1}$$

This implies: $1 + \frac{\gamma}{C_0}\frac{(1-\overline{\tau}_k)F(K_0,L_0)(1-\theta^m)}{\varphi^m} = \frac{1}{1-\tau_{k1}}$, or:

$$\tau_{k1} = \frac{\gamma(1-\overline{\tau}_k)F(K_0, L_0)(1-\theta^m)/C_0}{1 + \gamma(1-\overline{\tau}_k)(1-\theta^m)(1-\alpha(1-\beta))/C_0}$$

or

$$\tau_{k1} = \frac{\gamma(1-\overline{\tau}_k)(1-\theta^m)}{\frac{C_0}{F(K_0,L_0)} + \gamma(1-\overline{\tau}_k)(1-\theta^m)(1-\alpha(1-\beta))}$$

Also from the Foc's, assuming $\psi_t = 0 \ \forall \ t$, we get:

$$\frac{1-\alpha}{1-L_0} + \frac{\alpha\gamma(1-\overline{\tau}_k)(1-\theta^m)F_{l0}}{C_0\varphi^m} = \frac{\alpha}{C_0}\left\{1 + \frac{\gamma(1-\overline{\tau}_k)F(K_0,L_0)(1-\theta^m)}{C_0\varphi^m}\right\}F_{l0}$$

which simplifies to:

$$\frac{1-\alpha}{1-L_0} = \frac{\alpha}{C_0}\left\{1 + \gamma(1-\overline{\tau}_k)(1-\theta^m)\left[\frac{F(K_0,L_0)}{C_0} - 1\right]\right\}F_{l0}$$

or

$$\frac{C_0}{F_{l0}} = \frac{\alpha(1-L_0)}{1-\alpha}\left\{1 + \gamma(1-\overline{\tau}_k)(1-\theta^m)\left[\frac{F(K_0,L_0)}{C_0} - 1\right]\right\} > \frac{\alpha(1-L_0)}{1-\alpha}$$

Since $F_{l0} = \frac{F_0(1-\gamma)}{L_0}$ we have:

$$\frac{C_0}{F(K_0,L_0)} > \frac{\alpha(1-L_0)}{(1-\alpha)L_0}(1-\gamma)$$

Since $L_t \leq \overline{L}$ for some $\overline{L}$ (derived bellow), we have:

$$\frac{C_0}{F(K_0,L_0)} > \frac{\alpha(1-\overline{L})}{(1-\alpha)\overline{L}}(1-\gamma)$$

and then the result follows.

Auxiliary Result: In any CE, for all t $L_t \leq \overline{L} \equiv \frac{\alpha(1-\underline{\tau}_l)}{(1-\underline{\tau}_l)+(1-\alpha)\underline{\tau}_l}$.

In any CE, the shares are:

$$\varphi^i = (1-\beta)\left[\frac{\alpha(1-\tau_{k0})F_{k0}}{C_0} + \sum_t \frac{\alpha\beta^t}{C_t}\left[(1-\tau_{lt})F_{lt} + T_t\right]\right]$$

Market clearing implies:

$$\sum_t \frac{\alpha\beta^t}{C_t} F_{lt}[1-\tau_{lt}(1-L_t)] = \frac{1}{1-\beta} - \frac{\alpha(1-\tau_{k0})F_{k0}}{C_0} - \sum_t \frac{\alpha\beta^t}{C_t}\tau_{kt}K_tF_{kt} > \frac{1}{1-\beta}$$

Since in a CE we have $\frac{\alpha}{C_t}F_{lt}(1-\tau_{lt}) = \frac{1-\alpha}{1-L_t}$, it follows that:

$$\sum_t \beta^t \frac{1-\alpha}{1-L_t}\left[\frac{1-\tau_{lt}(1-L_t)}{1-\tau_{lt}}\right] > \frac{1}{1-\beta}$$

Again form the CE condition $\frac{\alpha}{C_t}F_{lt}(1-\tau_{lt}) = \frac{1-\alpha}{1-L_t}$, the Cobb-Douglas assumption implies:

$$\alpha(1-\gamma)k_t^{\gamma}(1-\tau_{lt}) = \frac{(1-\alpha)L_t^{\gamma}}{1-L_t}[K_t^{\gamma}L_t^{1-\gamma} - K_{t+1}]$$

Holding the intertemporal decisions fixed, we have that L_t is a decreasing function of labor taxes. Therefore for any t, labor is maximum when $\tau_{lt} = -\underline{\tau}_l$. Using this finding in the last inequality above we get:

$$\frac{1-\alpha}{1-\overline{L}}\left[\frac{1-\underline{\tau}_l(1-\overline{L})}{1-\underline{\tau}_l}\right]\frac{1}{1-\beta} > \frac{1}{1-\beta}$$

or $\overline{L} \equiv \frac{\alpha(1-\underline{\tau}_l)}{(1-\underline{\tau}_l)+(1-\alpha)\underline{\tau}_l}$ $\square$

17.12 Bounded Allocations

Here we show that, in any CE:

1. $\exists\ \underline{K}$ such that $\forall K_0 \in [\underline{K}, \overline{K}]$ and all feasible sequences of fiscal policy, $K_t \geq K_0\ \forall t$ in any CE.

2. Given $K_0 \in [\underline{K}, \overline{K}]$, in any CE we have $L_t \in [\underline{L}, \overline{L}]$ with $\underline{L} > 0$ and $\overline{L} < 1$.

In the proofs bellow, $F_k(K, L) = \gamma(L/K)^{1-\gamma}$ will denote the marginal product of capital.

<u>Lemma AA</u>:$\exists\ \underline{K}$ such that $\forall K_0 \in [\underline{K}, \overline{K}]$ and all feasible sequences of fiscal policy, $K_t \geq K_0\ \forall t$ in any CE.

Proof: If not, $\forall\ \underline{K_0} > 0\ \exists\ K_0 \geq \underline{K_0}$ and a sequence of feasible FP such that $K_t < \underline{K_0}$ for some t. Since the continuation of a CE is a CE, wlog assume t=1.

Hence consider a decreasing sequence $\{\underline{K_0^n}\}_{n=0}^{\infty}$ with $\lim_{n\to\infty} \underline{K_0^n} = 0$ and for all n there exists $K_0^n > \underline{K_0^n}$ and a sequence of feasible policies τ^n such that the CE of $\Gamma(K_0^n|\tau^n)$ has $K_1^n < K_0^n$.

<u>Fact 1</u>: For $t = 0$ or $t = 1$, there does not exist a subsequence such that $\lim_{ns\to\infty} L_t^{ns} =$

0 and $\lim_{ns\to\infty} L_t^{ns}/K_t^{ns} < \infty$.

<u>Proof of Fact 1</u>:If there exist such a sequence, then the marginal disutility of working for the mean type converges to $-(1-\alpha)$. Now consider an alternative plan at t: the mean type works a little and consumes immediately the labor income. The change in utility coming from increasing his consumption is $u_{c_t^{ns}} p_{l_t^{ns}} \equiv \frac{\alpha}{C_t^{ns}} \frac{(1-\gamma)(1-\tau_{lt}^{ns}) K_t^{ns\gamma}}{L_t^{ns\gamma}}$. If $\lim_{ns\to\infty} L_t^{ns} = 0$ this change goes grows arbitrarily large (hence it contradicts optimality for the mean type) unless $K_t^{ns} \to 0$. If, indeed, $K_t^{ns} \to 0$, then:

$$
\begin{aligned}
u_{c_t^{ns}} p_{l_t^{ns}} &= \frac{(1-\gamma)(1-\tau_{lt}^{ns})}{[K_t^{ns\gamma} L_t^{ns1-\gamma} + (1-\delta)K_t^{ns} - K_{t+1}^{ns}](L_t^{ns}/K_t^{ns})^{\gamma}} \\
&= \frac{(1-\gamma)(1-\tau_{lt}^{ns})}{L_t^{ns} + [(1-\delta)K_t^{ns} - K_{t+1}^{ns}](L_t^{ns}/K_t^{ns})^{\gamma}}
\end{aligned}
$$

Since $K_t^{ns} \to 0$, this Cauchy sequence is such that, for any $\epsilon > 0$, there exists N such that for all $ns \geq N$:

$$
\begin{aligned}
|(1-\delta)K_t^{ns} - K_{t+1}^{ns}| &= |(1-\delta)K_t^{ns} - (1-\delta)K_{t+1}^{ns} + (1-\delta)K_{t+1}^{ns} - K_{t+1}^{ns}| \\
&\leq (1-\delta)|(K_t^{ns} - K_{t+1}^{ns})| + \delta|K_{t+1}^{ns}| \\
&< (1-\delta)\epsilon + \delta\epsilon = \epsilon
\end{aligned}
$$

Therefore unless $\lim_{ns\to\infty} L_t^{ns}/K_t^{ns} = \infty$, the denominator of $u_{c_t^{ns}} p_{l_t^{ns}}$ above goes to zero. Either case it is a contradiction. $L_t^{ns}/K_t^{ns} \to \infty$ contradicts the initial assumption, while $u_{c_t^{ns}} p_{l_t^{ns}} \to \infty$ contradicts the optimality of the plan for the mean type. This finishes the proof of Fact 1.

Next, since $K_1^n \to 0$, $L_1^n/K_1^n \to \infty$. Otherwise $L_1^n/K_1^n \to L/K < \infty$ would imply $L_1^n \to 0$, contradicting Fact 1. Therefore $F_k(K_1^n, L_1^n) \to \infty$.

Fact 2: $\limsup_{n\to\infty} K_0^n = 0$.

Proof of Fact 2:We have two cases.

Case 1: $\liminf_{n\to\infty}(1 + F_k(K_0^n, L_0^n) - \delta)K_0^n > 0$.

Notice that this is the capital income of the median type after transfers in equilibrium. In this case C_0^n would be bounded away from zero and $F_k(K_0^n, L_0^n)$ becomes arbitrarily large. This means that the mean type could improve on the original plan by saving a little at period 0 and consuming the additional income in period 1. Notice since $C_0^n \geq \epsilon > 0$, the marginal utility lost in period 0 is bounded. This improvement on the original plan is a contradiction.

Case 2: $F_k(K_0^n, L_0^n) \to 0$. Since $K_0^n \to 0$, $F_k(K_0^n, L_0^n) \to 0$ is only possible if $L_0^n \to 0$. But using Fact 1 this would imply $F_k(K_0^n, L_0^n) \to \infty$. This finishes the proof of Fact 2.

Since $K_0^n \to 0$, $L_0^n/K_0^n \to \infty$ from Fact 1. Then we have:

$$\begin{aligned} C_0^n &\geq F_k(K_0^n, L_0^n)K_0^n + (1-\delta)K_0^n - K_1^n \\ &\geq F_k(K_0^n, L_0^n)K_0^n - K_1^n \\ &> (F_k(K_0^n, L_0^n) - 1)K_1^n \\ &\geq 0 \end{aligned}$$

where the strict inequality comes from $K_1^n < K_0^n$ and the last line comes from $F_k(K_0^n, L_0^n) \to \infty$. Since $F_k(K_0^n, L_0^n) \to \infty$ and consumption is bounded away from zero, again this leads to a contradiction: for n large enough the mean type would be better of by saving a little more and consuming the additional income in the next period ■

<u>Lemma BB</u>: Given $K_0 \in [\underline{K}, \overline{K}]$, there exists $(\underline{L}, \overline{L})$, potentially dependent on $\underline{K}$, such that any CE has $L_t \in [\underline{L}, \overline{L}]$ with $\underline{L} > 0$ and $\overline{L} < 1$.

<u>Proof</u>: We start by showing $L_t \geq \underline{L} > 0$. Towards a contradiction, suppose that there exists a subsequence $\{tn\}$ with $L_{tn} \to 0$.Using exactly the same reasoning as in the proof of Fact 1, it can be shown that $L_{tn} \to 0$ implies $L_{tn}/K_{tn} \to \infty$. But this is possible only if $K_{tn} \to 0$, which contradicts Lemma AA.

Next we show that $L_t \leq \overline{L} < 1$. Towards a contradiction, suppose that there exists a subsequence $\{tn\}$ with $L_{tn} \to 1$. Then the marginal disutility from labor goes to $-\infty$. The following alternative plan will improve his utility: for tn large enough, work a little less and the decrease the consumption by the amount in which his labor income diminishes. His change in consumption is $\frac{\alpha}{C_{tn}}(1-\gamma)(1-\tau_{ltn})(K_{tn}/L_{tn})^{\gamma} \rightarrow R < \infty$ unless $C_{tn} \rightarrow 0$. But indeed C_t is bounded away from zero because the log preferences and the fact that income from capital is bounded away from zero (from Lemma AA) ■