Pareto-Improving Optimal Capital and Labor Taxes*

Katharina Greulich† Albert Marcet‡

February 19, 2008

Abstract

We show a standard model where the optimal tax reform is to cut labor taxes and leave capital taxes very high in the short and medium run. Only in the very long run would capital taxes be zero. Our model is a version of Chamley's, with heterogeneous agents, without lump sum transfers, an upper bound on capital taxes, and a focus on Pareto improving plans. For our calibration labor taxes should be low for the first ten to twenty years, while capital taxes should be at their maximum. This policy ensures that all agents benefit from the tax reform and that capital grows quickly after when the reform begins. Therefore, the long run optimal tax mix is the opposite from the short and medium run tax mix. The initial labor tax cut is financed by deficits that lead to a positive long run level of government debt, reversing the standard prediction that government accumulates savings in models with optimal capital taxes. If labor supply is somewhat elastic benefits from tax reform are high and they can be shifted entirely to capitalists or workers by varying the length of the transition. With inelastic labor supply there is an increasing part of the equilibrium frontier, this means that the scope for benefitting the workers is limited and the total benefits from reforming taxes are much lower.

*We wish to thank Jordi Caballé, Begoña Domínguez, Joan M. Esteban, Andreu Mas-Colell, Michael Reiter, Sevi Rodríguez, Kjetil Storesletten, Jaume Ventura, Iván Werning, Philippe Weil, and Fabrizio Zilibotti for useful comments. M. Reiter provided the fixed point subroutine used to solve the model. Marcet acknowledges CREI, DGES, CIRIT (SGR2005-00097) and Barcelona GSE Research and the XREA network for their support. Marcet's address (corresponding author) albert.marcet@iae.csic.es, Institut d'Anàlisi Econòmica CSIC, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain. Greulich acknowledges support from the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK) and the Research Priority Program on Finance and Financial Markets of the University of Zurich. Greulich's address: greulich@iew.uzh.ch.

†IEW, University of Zurich

‡Institut d'Anàlisi Econòmica CSIC, CEPR

1 Introduction

A large literature on optimal dynamic taxation concludes that long run capital taxes should be zero. This result, which originally goes back to Chamley (1986) and Judd (1985), has been very resilient to many modifications of the basic model.[1]

That capital taxes should be so low is a controversial policy recommendation. Given the highly skewed distribution of wealth it would seem that lowering capital taxes and increasing labor taxes instead will necessarily hurt less wealthy taxpayers. But it is well known that in standard models capital taxes should be zero in the long run even with heterogeneous agents, and even if the government only considers policy allocations that improve agents with very little wealth. Chamley (1986), Judd (1985) and Atkeson, Chari & Kehoe (1999), provide results of this kind in different settings. In keeping with the literature we call this the Chamley/Judd result.

One interpretation of this result has been that there is no equity/efficiency trade-off involved in lowering capital taxes. It suggests that any opposition to lower capital taxes can only be due to a lack of understanding of economics, or to a belief in myopic behavior on the part of agents, or to inefficiencies of the political system that make it impossible for the government to commit, or to some other failure of the basic model. Consequently, many economists hold the view that introducing heterogeneity in models of optimal factor taxation with infinitely-lived agents is a nuisance.[2]

Upon closer inspection we find that optimal factor taxation depends very much on heterogeneity. If all agents have to benefit from a tax reform, optimal capital taxes should be high and, in our main model, labor taxes should be very low during a very long transition, between 10 and 20 years for our calibration. Only in the very long run should capital taxes be zero.[3]

[1] The literature is very large, a fair summary would be quite long. A very incomplete summary of the literature is that in the few cases where steady state optimal capital taxes are not zero they are often small or even negative.

[2] Some papers consider taxation in heterogeneous agent models with exogenous policy, we mention a few in the text. Optimal policy analysis in models with heterogeneous infinitely-lived agents is available, for example, in Bassetto (1999) and Niepelt (2002), who study how taxes affect taxpayers of different wealth in a stochastic model without capital; and Werning (2007) who studies redistribution with progressive taxation.

[3] Aiyagari (1995) showed that capital taxes should be positive in the *long run* due to capital overaccumulation in a model with heterogeneous agents and incomplete markets. We do not focus on this implication of heterogeneous agents for two reasons. First because the result is tenuous, Chamley (2001) shows that depending on the stochastic form of income shocks the long run capital tax should be negative; Marcet, Obiols-Homs &

One implication of these results is that long run properties of optimal policies should not be used for policy recommendations, since an optimal reform calls for many years of very low labor taxes and high capital taxes, the exact opposite of the long run recommendation.

Our results are complementary to some papers already hinting that the transition of optimal policy is very important in models of heterogeneous agents. These papers establish that in models of heterogeneous agents large parts of the population would suffer a large utility loss if capital taxes were suddenly abolished. Relevant references are Garcia-Milà, Marcet & Ventura (1995) (a model without uncertainty and calibration according to wage/wealth ratios), Correia (1995) (some analytic results), Domeij & Heathcote (2004) (a model with incomplete markets), Conesa & Krueger (2006) (with overlapping generations), and Flodén (2006) (policy designed optimally for one of the agents). The results of these contributions stand in stark contrast to Lucas (1990), who showed that the welfare of a representative agent would increase if capital taxes were abolished immediately and all tax revenue were obtained by taxing only labor. Thus, while designing the transition of capital/labor taxes optimally may not be very important with homogeneous agents, with heterogeneous agents there is indeed an important equity/efficiency trade-off and optimally designing the policy along the transition is crucial in resolving this issue.

Our results may also provide some insights for the political economy literature. Capital taxes in the real world are indeed very high, between 40% and 70% (after depreciation allowances) depending on the measurements. Having the Chamley/Judd result in mind this might seem like a failure of the institutions that, in the real world, determine fiscal policy. Indeed some papers in the political economy literature are able to explain high taxes as an outcome of frequent voting. One might conclude that low level of capital taxes should be part of an immutable constitution. But, in the light of our results, high capital taxes can be part of an optimal reform and they are not necessarily a failure of a political system. The puzzle now would become, why are labor taxes so high?.

Our results are further in line with the literature on gradualism of political reforms,

Weil (2007) argue that a similar result holds when introducing endogenous labor supply. Second, the result is specifically for the "veil of ignorance" welfare function, it may not hold for other welfare functions. Thus, because our aim is to study the transition, so we prefer to stay within a model where the zero long run optimal capital tax result holds.

which has been at the center of some policy debates.[4] The very long period of high capital taxes we find can be seen as a gradual reform designed to ensure that all agents improve their welfare.

We derive our results in a model very close to Chamley's benchmark, but with heterogeneous agents, without lump sum transfers, with an upper bound on capital taxes below 100%, and with a focus on Pareto improving plans. These features first of all create a meaningful equity-efficiency trade-off. Moreover, the restriction to Pareto improving allocations is natural because the surprising part of the Chamley/Judd result is that long run capital taxes should be zero even if the government improves all agents' welfare.[5] Aside from this literature-driven motivation, it seems that a sufficiently large and angry minority can block a tax reform, or it may credibly threaten to overturn the reform in a future vote, so that in order to change the taxation status quo a sufficiently large part of the population should agree. While keeping as close as possible to Chamley, we deviate from much of the optimal policy literature in explicitly studying the entire path of optimal capital and labor taxes, and not only the steady state.[6]

In our model optimal capital taxes are still zero in the long run, as capital taxes will be at the upper bound for N periods and then transit to zero in two periods. We find that redistributive concerns cause the transition to be very long, capital taxes are high for between ten and twenty years (for our calibration), depending on exactly which Pareto improving allocation is selected. This long period of high capital taxes is needed in order to raise more tax revenues from the "capitalists" and less from the "workers"; only then all agents benefit from the tax reform.

To demonstrate the effects of heterogeneity in isolation we first study a model with a completely inelastic labor supply. In this case the first best is achieved under homogeneous agents by setting capital taxes to zero in all periods. But with heterogeneous agents capital taxes should remain at their upper bound for a very long time before they are abolished.

[4] For example, the desirable speed of transition to market economies of formerly planned economies has been extensively discussed both in policy and academic circles. Within this literature, closest to our approach is Lau, Qian & Roland (2001) who show a gradual reform that improves all agents' welfare.

[5] The recent work of Flodén (2006) studies policies that are optimal only for one agent, thus Pareto improvements are not necessarily achieved.

[6] Some papers have studied the transition in models of optimal policy. For example Jones, Manuelli & Rossi (1993) study the transition in several homogeneous agent models.

The reason is that zero capital taxes in all periods would leave "workers" worse off than in the status quo. Therefore, even though the planner has access to non-distortive labor taxes he/she has to resort to distortive capital taxation to lower the workers' tax bill. The resulting total welfare losses are quite large, but they are needed to ensure a Pareto improvement.

If labor supply is somewhat elastic capital taxes should also be very high for about ten to twenty years before they are set to zero. But now we find that labor taxes should be *lower* than status quo during this transition. Lower initial labor taxes increase labor supply thus promoting growth in the early periods. With this policy the government achieves a lower relative welfare loss.

An interesting aspect of these results is that the short-medium run optimal policy (high capital taxes, low labor taxes) is exactly the opposite of the steady state (zero capital taxes, high labor taxes). Ignoring the transition would yield very low welfare for some agents, as found in the literature.[7] This suggests that it is important to go beyond steady state analysis in studies of optimal policy. Zero capital taxes in the long run are only Pareto optimal and Pareto improving if they go along with high capital taxes and low labor taxes during the transition.

We also report results for a welfare function that weighs equally both agents. This is the most popular welfare function in the literature, it can be justified by assuming the planner chooses under the "veil of ignorance" or with probabilistic voting and equal political power of both agents. It turns out that this policy is Pareto improving and that it redistributes most of the welfare gains to the worker. The transition period of high capital taxes and low labor taxes is very long, about 18 years.

In our main model government debt is positive in the long run. This is because the government initially runs a deficit to finance the initial drop in labor taxes. The behavior of long run debt is, therefore, the opposite from the standard case under capital taxation, where the government often accumulates savings. In the face of the recently renewed interest in studying the determinants of the optimal level of debt[8] this shows that a

[7] As shown in Garcia-Milà et al. (1995), Correia (1995), Domeij & Heathcote (2004), Conesa & Krueger (2006) and Flodén (2006).

[8] See Faraglia, Marcet & Scott (2006) and the references therein.

positive level of government debt can be a by-product of an optimal reform.

Our focus on Pareto improvements leads to an analysis of the utility frontier of equilibria. This frontier is decreasing and concave in models without distortions, but as is well known it may be non-well-behaved in models with proportional taxes. Interestingly, in the fixed labor supply case the frontier of equilibria has an increasing part, even in the range of equilibria that dominate the status quo, and the government can only leave the capitalist in the status quo by pursuing a Pareto inefficient policy.

Many of our results are obtained numerically. To solve the model we need to take care of some technical issues, including the fact that the heterogeneity parameters (the choice variable λ defined below) have to be solved for separately, that the frontier of equilibria may not be well-behaved, and the non-recursiveness on the solution induced by the tax limit.

The main results are robust to various model and parameter changes. We also explore if progressive taxation might achieve the proper redistribution to ensure Pareto improvements and avoid the distortions associated with high capital taxes. Finally, we also numerically investigate the time consistency of our solutions. We find that the tax reform is time consistent if it can only be overturned by consensus, therefore, heterogeneity builds in some time-consistency.[9] However, if the planner reoptimizes in some future date using a welfare function with fixed weights the policy is time inconsistent.

The paper is organized as follows: In section 2 we lay out our baseline model and discuss further the motivation for our assumptions. Section 3 discusses some properties of the models obtained analytically, including a proof that capital taxes are zero in steady state and about the form of the transition. We calibrate the heterogeneity parameters of the two agents to the groups of highest and lowest 20% wealth/wage ratios. Our numerical results are discussed in section 4. Section 5 concludes.

2 The Model

We consider a model with two heterogeneous agents, discrete time, no uncertainty, Ramsey equilibrium, capital accumulation. This is almost a special case of Chamley (1986), our

[9]This is a similar result to the one obtained analytically by Armenter (2004) for a simpler model.

emphasis changes in four aspects relative to most of the literature:

i) We study the whole path for taxes, including the transition.

ii) We preclude agent-specific redistributive lump sum transfers. It is well known that the Chamley/Judd result survives even in this case.[10] This assumption seems reasonable in a literature that has focussed on the effects of distorting taxation and because most tax codes and indeed most constitutions stipulate that all individuals are equal in front of the law.

iii) We search for allocations that improve everybody in the population. We therefore calibrate the features of our two agent groups such as to represent the two extreme quintiles of the population and study plans that are Pareto improvements under this calibration. We also study the (veil of ignorance) case where the planner has a welfare function that weighs each agent type according to the population share of each type.

iv) We impose an upper bound on capital taxes each period. Chamley (1986) and Atkeson, Chari and Kehoe (1999) assume an upper bound of 100% for capital taxes in all periods. Many other papers in the optimal taxation literature tend to simplify things by assuming a bound only in the initial period. Optimal policies under these constraints imply that capital taxes should be very high in the first few periods, much higher than current actual capital taxes which, by all measures, are already very high. The initial tax hike recommended by these models could have devastating effects on investment in the real world if there is partial credibility of government policy, or if agents form their expectations by learning from past experience.[11] So, to avoid this tax hike in the initial periods we use capital tax ceilings lower than 100%. In all of our computational exercises we fix this ceiling to the status quo capital tax. Alternatively, this bound can be interpreted as the value that avoids massive capital flight in an open economy with partial mobility of

[10] To our knowledge the first complete proof that capital taxes are zero in the long run even in the absence of agent-specific lump sum redistribution is in section B1 of Chari & Kehoe (1999) also described in Atkeson, Chari and Kehoe (1999). Chamley (1986) discusses the case of heterogeneous agents by considering a government with a welfare function that weighs all agents linearly, this implicitly assumes there are redistributive lump sum transfers. Judd (1985) considered a model where an agent has only capital income and another agent has only wage income, so the distorting taxes in that paper are, in a way, agent-specific.

[11] Lucas (1990) offered a similar reasoning to motivate his study of a tax reform that abolishes capital taxes immediately. Of course, one could infer from our discussion that issues such as credibility and learning should be introduced explicitly in the analysis, instead of indirectly with the tax limit. Doing so would imply deviating very much from Chamley's model, but in this paper we prefer to stay as close as possible to that model in order to understand the reason for our results. The time consistency literature deals, in a way, with the credibility issue. An analysis of capital taxes under learning can be found in Giannitsarou (2006).

capital.

In the following we present our model formally. We refer to Garcia-Milà et al. (1995) for some details on how to characterize competitive equilibria. For details on formulating Ramsey equilibria and the primal approach in general, see Chari & Kehoe (1999) or Ljunqvist and Sargent (2002).

2.1 The environment

There are two consumers $j = 1, 2$ with utility $\sum_{t=0}^{\infty} \delta^t \left[u(c_{j,t}) + v(l_{j,t})\right]$ where c is consumption and l is labor of each agent each period. We assume $u' > 0$, $v' < 0$ and the usual Inada and concavity conditions. Agents differ in their initial wealth $k_{j,-1}$ and their labor productivity ϕ_j. Agent j obtains income in period t from renting his/her capital at the rental price r_t and from selling his/her labor for a wage $w_t\phi_j$. Agents pay taxes at rates τ_t^l on labor income and τ_t^k on capital income net of depreciation allowances. Therefore period-t budget constraint of agent j is given by

$$c_{j,t} + k_{j,t-1} = w_t \ \phi_j \ l_{j,t}(1 - \tau_t^l) + k_{j,t-1}\left[1 + (r_t - d)(1 - \tau_t^k)\right] \qquad \text{for } j = 1, 2 \qquad (1)$$

Firms maximize profits, have a production function $F(k_{t-1}, e_t)$ where e is total efficiency units of labor and k is total capital, F is concave, increasing in both arguments, has constant returns to scale, $F_k(k, e) \to 0$ as $k \to \infty$, $F_{kk}(k, e) < 0$ for all $e > 0$ and $F_{ee}(k, e) < 0$ for all $k > 0$.

The government chooses capital and labor taxes, consumes g in every period, and has the standard budget constraint. It saves in capital and has initial capital k_{-1}^g. Government can get in debt, that is, k^g can be negative. Ponzi schemes for consumers and government are ruled out.

We normalize each agents' mass to be 1/2. Market clearing conditions are

$$\frac{1}{2}\sum_{j=1}^{2}\phi_j l_{j,t} = e_t \tag{2}$$

$$k_t = k_t^g + \frac{1}{2}\sum_{j=1}^{2} k_{j,t}$$

$$\frac{1}{2}\sum_{j=1}^{2} c_{j,t} + g + k_t - (1-d)k_{t-1} = F(k_{t-1}, e_t) \tag{3}$$

2.2 Competitive Equilibrium Conditions

The equilibrium concept is standard: consumers take prices and taxes as given, they maximize their own utility, markets clear, and the budget constraint of the government is satisfied. The first order conditions of optimality with respect to capital and labor for consumer j are

$$u'(c_{j,t}) = \delta\ u'(c_{j,t+1})\ \left(1 + (r_{t+1} - d)(1 - \tau_{t+1}^k)\right) \tag{4}$$

$$-\frac{v'(l_{j,t})}{u'(c_{j,t})} = w_t\ (1 - \tau_t^l)\ \phi_j \tag{5}$$

for all t and j. Firms' profit maximization implies that factor prices equal marginal product: $r_t = F_k(k_{t-1},\ e_t)$ and $w_t = F_e(k_{t-1},\ e_t)$.

We also assume

$$u(c) = \frac{c^{1-\sigma_c}}{1-\sigma_c} \qquad v(l) = B\frac{(1-l)^{1-\sigma_l}}{1-\sigma_l} \tag{6}$$

In this case FOC for capital and labor imply

$$\frac{c_{2,t}}{c_{1,t}} = \lambda \ \text{ and } \ \frac{1-l_{2,t}}{1-l_{1,t}} = \lambda^{\frac{\sigma_c}{\sigma_l}}\left(\frac{\phi_2}{\phi_1}\right)^{\frac{1}{\sigma_l}} \qquad \text{for all } t \tag{7}$$

for some λ that is constant through time.

Using (7) and standard steps of the primal approach one can show that (4), (5) and the budget constraints of the consumer j for all $t = 0, 1, ...$ can be summarized in the present value constraint[12]

$$\sum_{t=0}^{\infty} \delta^t \frac{u'(c_{1,t})}{u'(c_{1,0})} \left(c_{j,t} - w_t \ \phi_j \ l_{j,t}(1-\tau_t^l)\right) = \tag{8}$$

$$= k_{j,-1}(1+(r_0-d)(1-\tau_0^k)) \qquad \text{for} \quad j=1,2$$

Using (7) again one can show that necessary and sufficient conditions for a sequence $\{c_t^1, k_t, l_t^1\}$ and a constant λ to be a competitive equilibrium are

$$\sum_{t=0}^{\infty} \delta^t \left(u'(c_{1,t})c_{1,t} + v'(l_{1,t}) \ l_{1,t}\right) = u'(c_{1,0}) \ k_{1,-1}(1+(r_0-d)(1-\tau_0^k)) \tag{9}$$

$$\sum_{t=0}^{\infty} \delta^t \left(u'(c_{1,t})\lambda c_{1,t} + \frac{\phi_2}{\phi_1} v'(l_{1,t}) \ f(\lambda, l_{1,t})\right) = u'(c_{1,0}) \ k_{2,-1}(1+(r_0-d)(1-\tau_0^k)) \tag{10}$$

and feasibility (3). Here $f(\lambda, l_{1,t})$ is defined as

$$f(\lambda, l_{1,t}) \equiv 1 - (1-l_{1,t}) \, \lambda^{\frac{\sigma_c}{\sigma_l}} \left(\frac{\phi_2}{\phi_1}\right)^{\frac{1}{\sigma_l}} \tag{11}$$

and it gives $l_{2,t}$ that solves (7) for each possible value of the endogenous variables λ, l_t^1. It is possible to show that as long as these conditions hold capital and labor taxes can be found that ensure all first order conditions of agents hold. See Garcia-Milà et al. (1995) for details. Notice that condition (9) is standard for agent 1 in models of optimal policy but that (10) is an analogous condition for agent 2 when equilibrium conditions (7) are imposed. Note that the presence of heterogeneous agents implies that the ratio λ has to be found optimally subject to the constraints (9) and (10).

Taxes are then found as a residual from (4) and (5). Consumption and labor of agent 2 are found from (7) and individual capital is backed out from the budget constraint period by period.

2.3 Policy Objective and Constraints

We assume that the planner chooses *Pareto optimal* allocations. A standard argument justifies that this is equivalent to assuming that the planner maximizes the utility of, say,

[12] Walras' law guarantees that the budget constraint of the government is implied by the above equations plus feasibility so it can be ignored.

agent 1, subject to the constraint that agent 2 utility has a minimum value $\underline{U}^2$:

$$\sum_{t=0}^{\infty} \delta^t \left[u(c_{2,t}) + v(l_{2,t})\right] \geq \underline{U}^2 \tag{12}$$

where $\underline{U}^2$ is restricted so that the set of feasible competitive equilibria satisfying this constraint is non-empty. Varying the value of the minimum utility $\underline{U}^2$ along all possible utilities that can be achieved in equilibrium for agent 2 we can trace out the whole set of Pareto efficient allocations.

We will concentrate our attention on allocations that are *Pareto improving* relative to a certain status quo. Let U^j_{SQ} be the status quo utility obtained by agent j, achieved with some taxation scheme that is already in place.[13] Clearly these can be found by considering only minimum utility values such that $\underline{U}^2 \geq U^2_{SQ}$ *and* such that the planner's objective at the maximum satisfies

$$\sum_{t=0}^{\infty} \delta^t \left[u(c^*_{1,t}) + v(l^*_{1,t})\right] \geq U^1_{SQ}$$

where * denotes the optimized value of each variable for a given $\underline{U}^2$. We refer to these Pareto optimal and Pareto improving plans as "POPI" allocations. Proposition 2 below will provide a way to compute all the utility values on the frontier and to select the POPI allocations.

Finally, we introduce *tax limits* $\tau^k_t \leq \widetilde{\tau}$ for all $t = 0, 1, ...$, ensuring that capital taxes never go above a certain constant $\widetilde{\tau}$ exogenously given. Combining this limit with the FOC for capital of the consumers it is easy to see that the tax limit is satisfied in equilibrium if and only if

$$u'(c_{1,t}) \geq \delta\ u'(c_{1,t+1})\ \left(1 + (r_{t+1} - d)(1 - \widetilde{\tau})\right) \quad \text{for all } t > 0 \quad \text{and} \tag{13}$$

$$\tau^k_0 \leq \widetilde{\tau} \tag{14}$$

The first equation ensures that the actual capital tax τ^k_t for $t = 1, 2, ...$ that is implied by (4) satisfies the limit and it allows us to use use the primal approach where taxes do not

[13] The status quo utility, in general, depends on the distribution of capital at period -1, but we leave this dependence implicit.

appear explicitly in the constraints of the government.

2.4 The Lagrangian

As standard in the literature we look for a Ramsey equilibrium where the government chooses an optimal sequence of tax rates and deficits, maximizes utility of agent 1 subject to the constraint that taxes and prices have to be compatible with competitive equilibrium and subject to the above additional constraints. The government has full credibility, full commitment to the preannounced policy, both government and agents have rational expectations. So, the government/planner solves

$$\begin{aligned} \max_{\tau_0^k, \lambda, \{c_t^1, k_t, l_t^1\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \delta^t \; [u(c_{1,t}) + v(l_{1,t})] \\ \text{s.t.} \quad & \sum_{t=0}^{\infty} \delta^t \; [u(\lambda c_{1,t}) + v(f(\lambda, l_{1,t}))] \geq \underline{U}^2 \end{aligned} \tag{15}$$

and subject to feasibility (3) for all t, the implementability constraints (9) and (10) (for period 0 only) and the tax limits (13), (14). Here, we used (7) and (11) to substitute equilibrium c_2, l_2 out in (15). $\underline{U}^2$ has to satisfy the requirements discussed in the previous subsection to achieve a Pareto improvement.

Notice that a special feature of this problem is that the constant λ has to be determined as a feature of the optimal choice, therefore it appears as an argument in the maximization problem.

Let α be the Lagrange multiplier of the minimum utility constraint (15), let Δ_1, Δ_2 be the multipliers of (9) and (10) normalized by $u'(c_{1,0})$, and let γ_t be the Lagrange multiplier of (13), the Lagrangian for the government problem is

$$
\begin{aligned}
L = \sum_{t=0}^{\infty} \delta^t \Big[\; & u(c_{1,t}) + v(l_{1,t}) + \alpha \left[u(\lambda c_{1,t}) + v(f(\lambda, l_{1,t})) \; \right] + \\
& \Delta_1 [\; u'(c_{1,t}) c_{1,t} + v'(l_{1,t}) \; l_{1,t} \;] + \\
& \Delta_2 [\; u'(c_{1,t}) \lambda c_{1,t} + \frac{\phi_2}{\phi_1} v'(l_{1,t}) f(\lambda, l_{1,t}) \;] + \\
& \gamma_t [\; u'(c_{1,t}) - \delta \; u'(c_{1,t+1})(1 + (r_{t+1} - d)(1 - \tilde{\tau})) \;] - \\
& \mu_t \left(\frac{1+\lambda}{2} c_{1,t} + g + k_t - (1-d) k_{t-1} - F(k_{t-1}, e_t) \right) \; \Big] - \mathbf{W} - \alpha \underline{U}^2
\end{aligned} \tag{16}
$$

where $\mathbf{W} = u'(c_{1,0}) \left[[\Delta_1 k_{1,-1} + \Delta_2 k_{2,-1}] \left[1 + (r_0 - d)(1 - \tau_0^k) \right] \right]$. Further, $\gamma_t, \alpha \geq 0$ and they satisfy the usual slackness conditions.

The first line of this Lagrangian has the usual interpretation: a Pareto efficient allocation amounts to maximizing a welfare function where the planner weighs linearly the utility of both agents, where the weight of agent 1 is normalized to one and the weight of agent 2 is the Lagrange multiplier of the minimum utility constraint. However, it is important that this α is not chosen arbitrarily in our setup, it has to be found to satisfy the Pareto improving constraint (15). The next two lines ensure the budget constraints of the consumers, the fourth line is the upper bound on capital taxes and the last line is the feasibility constraint. The term $\mathbf{W}$ collects the period 0 terms in the budget constraints of the consumers.

As is often the case in optimal taxation models the feasible set of sequences for the planner is non-convex. This means that we will need to be careful about necessity and sufficiency of first order conditions. We will be explicit about these issues in section 3.2. First order conditions with respect to capital, labor and consumption are derived in a standard way and they are shown in appendix A. In the rest of the section we comment on features of these first order conditions that differ from other papers on dynamic taxation.

Since the relative consumption of agents λ is a choice we need to set the derivative of

L with respect to λ equal to zero, to obtain

$$\sum_{t=0}^{\infty} \delta^t \Big[\alpha \; [\; u'(\lambda c_{1,t}) \; c_{1,t} + v'(f(\lambda, l_{1,t})) \; f_\lambda(\lambda, l_{1,t})) \;] + \\ \Delta_2 [\; u'(c_{1,t}) \; c_{1,t} + \frac{\phi_2}{\phi_1} v'(l_{1,t}) \; f_\lambda(\lambda, l_{1,t}) \;] - \\ \frac{\mu_t}{2} [\; c_{1,t} - F_e(k_{t-1}, e_t) \; \phi_2 \; f_\lambda(\lambda, l_{1,t}) \;] \Big] = 0 \tag{17}$$

The fact that λ has to be chosen is a reflection of the fact that the government can vary the ratio of consumptions of the agents by varying the total tax burden of labor or capital in discounted present value.

The multipliers $\gamma, \Delta_1, \Delta_2$ have to satisfy the complementary slackness conditions. For α (the multiplier of (15))

$$\text{either } \alpha > 0 \text{ and } \sum_{t=0}^{\infty} \delta^t \; [u(c_{2,t}) + v(l_{2,t})] = \underline{U}^2$$

$$\text{or } \alpha = 0 \text{ and } \sum_{t=0}^{\infty} \delta^t \; [u(c_{2,t}) + v(l_{2,t})] \geq \underline{U}^2$$

In other words, the minimum utility constraint may or may not be binding. In the first case the Lagrangian amounts to maximizing the weighted utility of agents 1 and 2 with weight 1 and α, respectively. If the minimum utility constraint is NOT binding the planner gives zero weight to agent 2. The latter case would only occur in models without frictions if the planner would be willing to give a very low utility to agent 2, but we will see that it occurs in our case even if the lower bound $\underline{U}^2$ is the status quo utility. This is because even if $\alpha = 0$ agent 2 will be consuming due to the fact that the allocations are determined in equilibrium and the budget constraint of agent 2 has to be satisfied, insuring agent 2 some revenue for any policy action.

Similarly, for the γ's and for each t, we have

$$\text{either } \gamma_t > 0 \text{ and } u'(c_{1,t}) = \delta \; u'(c_{1,t+1})(1 + (r_{t+1} - d)(1 - \tilde{\tau}))$$

$$\text{or } \gamma_t = 0 \text{ and } u'(c_{1,t}) \geq \delta \; u'(c_{1,t+1})(1 + (r_{t+1} - d)(1 - \tilde{\tau}))$$

It turns out that the Δ_i's may be positive or negative, since the corresponding present value budget constraints have to be satisfied with equality. This becomes clear by looking at the following interpretation. With two agents the marginal utility cost of distortive taxation is $\frac{\partial L}{\partial \tau_0^k} = u'(c_{1,0})[\Delta_1 k_{1,-1} + \Delta_2 k_{2,-1}](r_0 - d)$. Thus,

$$\Delta_1 k_{1,-1} + \Delta_2 k_{2,-1} \geq 0$$

with the inequality being strict as long as any taxes are raised after the initial period. This allows for one of the Δ_i being negative, which will indeed be the case whenever the constraints on redistribution that are imposed by the competitive equilibrium conditions and the Pareto improvement requirement are sufficiently severe. To see this consider a slightly modified model in which the social planner is allowed to redistribute initial wealth between agents by means of a lump sum transfer T^i such that $T^1 = -T^2$. All this modification does to the Lagrangian is to change the implementability constraints such that the term $(\Delta_1 - \Delta_2)T^1$ is added inside the large bracket in the definition of $\mathbf{W}$. Now the derivative of the Lagrangian with respect to the lump-sum transfer between agents is $\frac{\partial L}{\partial T^1} = u'(c_{1,0})(\Delta_1 - \Delta_2)$. For any given T^1, and in particular for $T^1 = 0$ as in our baseline model, this expression is a measure of the marginal utility cost of the transfer not being optimal. If the planner were free to choose T^1 optimally, we would have $\Delta_1 = \Delta_2 > 0$. If the planner would like to redistribute more towards agent 2, $\Delta_1 - \Delta_2 > 0$ and vice versa. If the transfer is much too low (high) the derivative will be large in absolute value and Δ_2 (Δ_1) will be negative. In sum, while the weighted sum of the multipliers on the present value budget constraints is related to the cost of distortive taxation, their difference indicates the cost of not being able to redistribute lump sum. These multipliers thus capture in a simple way the two forces that drive the solution to our model away from the first best: the absence of lump-sum taxes and of agent-specific lump-sum transfers.

Throughout the paper we will consider variants of this basic model that we will dub "modified models". These will be useful to obtain some analytic results and to demonstrate the behavior of the model.

For the government problem to be well defined we should ensure that the set of feasible equilibria is non-empty. This is guaranteed, for example, by the existence of a status quo equilibrium, if $\widetilde{\tau}$ is larger or equal to the status quo capital tax, and if $\underline{U}^2$ is close to the status quo utility.

3 Characterization of equilibria

Here we describe some analytical results.

3.1 Qualitative behavior of capital taxes

First of all we derive the behavior of the economy in steady state and we describe some properties of the transition. To the best of our knowledge there is no previous proof of zero long-run capital taxes that fully applies to our model.[14]

To obtain this result we assume the government has free disposal of g. More precisely, we assume that the government can purchase consumption good in excess of g and dump the excess. It is easy to see that this is equivalent to assuming that the feasibility constraint (3) holds as an inequality $\leq$ instead of as equality.

Proposition 1: *Assume free disposal of g, log utility of consumption ($\sigma_c = 1$) and 1 $>\widetilde{\tau} > 0$. Assume that $F(k, 0) = F(0, e) = 0$. Assume the economy converges to a steady state such that $\overline{c} > 0$. Then the optimal capital tax rate jumps from the tax limit to zero in two periods. Formally, there is a finite N such that*

$$\begin{aligned} \tau_t^k &= \tilde{\tau} \quad \textit{for all } t \leq N, \\ &= 0 \quad \textit{for all } t \geq N+2, \end{aligned}$$

Proof:

We proceed in two steps. First we show that it is not possible for the tax limit to be binding forever in the optimal allocation. Then we show that capital taxes go from the limit to zero in two periods.

[14]The results in Chari & Kehoe (1999) and Atkeson et al. (1999) are similar, they also prove the tax limit cannot be binding forever and that the transition takes two periods. But the results in those papers are not directly applicable here. They do not consider a tax limit and heterogeneity at the same time but, more importantly, their proof is for the case of a capital tax limit of 100%. For this particular bound if the tax limit were binding forever feasibility would be violated. In our case, where $\widetilde{\tau}$ is the status quo tax the same line of argument cannot be used: indeed the economy could stay at status quo forever. This is why a more involved argument is needed.

First of all, notice that in the log case the first order condition with respect to consumption for $t > 0$ becomes (see Appendix A)

$$\begin{aligned} c_{1,t}^{-1}\big(1+\alpha\lambda\big)- \\ -c_{1,t}^{-2}\big(\gamma_t - \delta\gamma_{t-1}(1+(F_k(k_{t-1},e_t)-d)(1-\tilde{\tau}))\big) = \mu_t\frac{1+\lambda}{2} \end{aligned} \tag{18}$$

Now we prove that the tax limit cannot be binding in all periods. If indeed the solution had $\tau_t^k = \tilde{\tau}$ for all t, then, denoting steady state variables with an upper bar,

$$\delta\left[1+(F_k(\overline{k},\overline{e})-d)(1-\tilde{\tau})\right] = 1 \quad \text{for all } t.$$

Evaluating (18) at steady state and plugging the last equation into the one above we have

$$\mathbf{A} + \overline{c}_1^{-2}\big(\gamma_t - \gamma_{t-1}\big) = \mu_t\frac{1+\lambda}{2} \tag{19}$$

where $\mathbf{A} = \overline{c}_1^{-1}(1+\alpha\lambda)$. The FOC for labor for $t > 0$ (see appendix A) at steady state implies

$$\begin{aligned} -B(1-\bar{l}_1)^{-\sigma_l}\big(1+\frac{\alpha}{\lambda}\frac{\phi_2}{\phi_1}f'(\lambda,\bar{l}_1)+\Delta_1+\Delta_2\frac{\phi_2}{\phi_1}f'(\lambda,\bar{l}_1)\big)- \\ \sigma_l B(1-\bar{l}_1)^{-\sigma_l-1}\big(\Delta_1\bar{l}_1+\Delta_2\frac{\phi_2}{\phi_1}f(\lambda,\bar{l}_1)\big)+ \\ \gamma_{t-1}\overline{c}_1^{-1}(1-\tilde{\tau})F_{k,e}(\overline{k},\overline{e})\frac{1}{2}(\phi_1+\phi_2 f'(\lambda,\bar{l}_1)) = \\ -F_e(\overline{k},\overline{e})\frac{1}{2}(\phi_1+\phi_2 f'(\lambda,\bar{l}_1))\mu_t \end{aligned} \tag{20}$$

Notice that we are only imposing steady state on the variables, not on the multipliers. This is the right way to proceed because real variables have natural bounds and existence of a steady state may be expected. But the multipliers should not have bounds, otherwise there is no sense in which the Lagrangian is guaranteed to give a maximum, and a steady state in the variables could be compatible with multipliers that go to infinity.

Collecting all the terms (20) that do not depend on the multipliers γ or μ we have

$$\mathbf{B} + \mathbf{C}\gamma_{t-1} = -\mu_t \tag{21}$$

for

$$\mathbf{B}=\Big[-B(1-\bar{l}_1)^{-\sigma_l}\big(1+\frac{\alpha}{\lambda}\frac{\phi_2}{\phi_1}f'(\lambda,\bar{l}_1)+\Delta_1+\Delta_2\frac{\phi_2}{\phi_1}f'(\lambda,\bar{l}_1)\big)-$$
$$\sigma_l B(1-\bar{l}_1)^{-\sigma_l-1}\big(\Delta_1\bar{l}_1+\Delta_2\frac{\phi_2}{\phi_1}f(\lambda,\bar{l}_1)\big)\Big]\frac{1}{F_e(\overline{k},\overline{e})\frac{1}{2}(\phi_1+\phi_2 f'(\lambda,\bar{l}_1))}$$
$$\mathbf{C}=\frac{\overline{c}_1^{-1}(1-\tilde{\tau})F_{ke}(\overline{k},\overline{e})}{F_e(\overline{k},\overline{e})}$$

Notice that the terms $\mathbf{A},\mathbf{B},\mathbf{C}$ do not depend on the multipliers γ or μ. So, combining (19) and (21) we have that near the steady state the evolution of γ is approximately given by

$$\gamma_t=\overline{c}_1^2\left[\mathbf{A}+\mathbf{B}\frac{1+\lambda}{2}\right]+\left[1+\mathbf{C}\overline{c}_1^2\frac{1+\lambda}{2}\right]\gamma_{t-1} \quad . \tag{22}$$

Since the steady state involves $\overline{c}>0$ assumptions $F(k,0)=F(0,e)=0$ imply that $\overline{k},\overline{e}>0$. This implies that $0<F_e(\overline{k},\overline{e})<\infty$. The constant returns to scale assumption and concavity implies that $F_{ke}(\overline{k},\overline{e})\overline{k}=-F_{ee}(\overline{k},\overline{e})\overline{e}>0$. All these observations imply

$$\frac{F_{ke}(\overline{k},\overline{e})}{F_e(\overline{k},\overline{e})}>0 \tag{23}$$

Since the assumption $\widetilde{\tau}<1$ and (23) imply $\mathbf{C}>0$, the last bracket in (22) is larger than one, and this is an unstable difference equation in γ_t. This implies that if the tax limit were binding at the steady state, $\gamma_t\to\infty$ or $-\infty$ depending on whether $\mathbf{A}+\mathbf{B}\frac{1+\lambda}{2}$ is positive or negative. It is impossible that $\gamma_t\to-\infty$, as the multiplier γ_t is non-negative. Also, $\gamma_t\to\infty$ implies, through (19), that μ_t is eventually negative, which is also impossible under the g-free-disposal assumption. Thus, the tax limit cannot be binding in all periods, there has to be a period t where $\tau_t^k<\widetilde{\tau}$.

Now we show that capital taxes go from the limit to zero in two periods. The previous argument implies that there is a finite $N+1$ which is the *first* period where the tax limit is not binding, so that $\tau_{N+1}^k<\widetilde{\tau}$ and $\tau_t^k=\widetilde{\tau}$ for all $t\le N$ in the optimum. Given N, consider the following modification to the baseline model. Assume that instead of the uniform tax limit in all periods we had considered a model where the tax limits are

$$\tau_t^k\le\widetilde{\tau}\quad\text{for all } t\neq N+1$$

but τ^k_{N+1} is unconstrained. Let us call this the "modified model 1" (MM1). It is clear that the solution to this problem is *equal* to the solution of the baseline model, because we have just relaxed a tax limit that was not binding in the optimum of the baseline model. Let us keep this fact in store for a while.

Now consider a second modified model, one that we dub MM2, where tax limits are

$$\tau^k_t \leq \widetilde{\tau} \quad \text{for all } t \leq N$$

but τ^k_t is unconstrained for all $t > N$. Let us denote with a $\widehat{\ }$ the solution to MM2.

Clearly the first order conditions for this modified model are the same as for the basic problem except that

$$\widehat{\gamma}_t = 0 \quad \text{for all } t \geq N \tag{24}$$

(Notice that γ_t is the multiplier associated with the constraint on τ^k_{t+1}, so that τ^k_{N+1} being the first unconstrained tax means γ_N is the first multiplier that must be 0.)

Combining (24) with (18), implies[15]

$$\widehat{c}^{-1}_{1,t}\left(1+\widehat{\alpha}\widehat{\lambda}\right) = \widehat{\mu}_t \frac{1+\widehat{\lambda}}{2} \quad \text{for all } t \geq N+1 \tag{25}$$

This last equation does not hold for $t = N$ because $\widehat{\gamma}_{N-1} \neq 0$ appears in (18). Plugging (24) in the FOC with respect to capital (see appendix A) we get

$$\widehat{\mu}_t = \delta\widehat{\mu}_{t+1}(1 + F_k(k_t, e_{t+1}) - d) \;\; \text{for all } t \geq N$$

and using (25) we have

$$\widehat{c}^{-1}_{1,t} = \delta\widehat{c}^{-1}_{1,t+1}(1 + F_k(k_t, e_{t+1}) - d) \;\; \text{for all } t \geq N+1$$

[15]Notice that in order to obtain the following equation we absolutely need log-utility. This equation would not hold for higher risk aversions because the term u'' in the FOC for consumption would not disappear in that case, so in that case and if capital is not exactly at steady state this does not obtain. Therefore, log utility is necessary in order obtain the proof. At this writing we are not sure why previous results on the transition of capital taxes with an upper bound did not incorporate log utility as an assumption.

Using the Euler equation of the consumer we conclude that

$$\widehat{\tau}_t^k = 0 \text{ for all } t \geq N+2 \quad . \tag{26}$$

Therefore, the properties for taxes mentioned in the statement of the proposition hold for the model MM2.

Since the optimal solution for MM2 is also feasible in MM1, even though the latter is in principle more restrictive because τ_t^k for $t > N+1$ are (potentially) constrained, $\widehat{\tau}_t^k$ is also the optimal tax in MM1. This proves that in MM1

$$\tau_t^k = 0 \text{ for all } t \geq N+1 \quad .$$

Since we already argued that the solution to MM1 was equal to the solution of the baseline model, this completes the proof.[16]■

The tax limit is a forward looking constraint and, therefore, standard dynamic programming does not apply. Using a promised utility approach would be complicated because of the appearance of a state variable (marginal utility of consumption) that has to be bounded to stay in the set of feasible marginal utilities and since there is a natural state variable k characterizing this set is quite difficult. The Lagrangian approach of Marcet & Marimon (1998) is easier to use in these circumstances. The details are worked out in Appendix C.

3.2 The Frontier of the Equilibrium Set

We now study the frontier of equilibrium utilities. Formally, let $\mathcal{F}$ be the frontier of the set

$$\Big\{(U_1, U_2) \in R^2 : U_i = \sum_{t=0}^{\infty} \delta^t \, [u(c_{i,t}) + v(l_{i,t})] \text{ for some } \{(c_{i,t}, l_{i,t})_{i=1,2}, k_t\} \text{ a CE}\Big\} \tag{27}$$

[16]Notice that for the proof to work we do need to consider the two modified models MM1 and MM2. If we tried to compare MM2 to the solution of the baseline model directly we would not be able to rule out that $\widehat{\tau}_{N+1}^k > \widetilde{\tau}$, the solution to MM2 would then be unfeasible in the baseline model and could not be compared to it. Considering MM2 as a restricted version of MM1 allows us to rule out $\widehat{\tau}_{N+1}^k > \widetilde{\tau}$.

In the standard case without distortions and a concave utility function it is well known that $\mathcal{F}$ coincides with the Pareto optimal frontier and it defines a decreasing function of, say, U_1 as a function of U_2. In this case all these allocations can be traced by optimizing welfare functions that give different weights to each agent. Given the distortions introduced by proportional taxes we cannot be sure that the set (27) is convex. Once we leave the beautiful world of maximizing concave objective functions over convex sets we are no longer sure if considering a welfare function allows us to trace the frontier of equilibria. Furthermore now the frontier of equilibria may not coincide with the set of Pareto optimal allocations. There are indeed models where the frontier of equilibria has a convex part that cannot be found by maximizing welfare functions.

But the situation is not totally desperate. We can still find sufficient conditions guaranteeing that by maximizing a welfare function in the standard way we obtain points in the frontier, and we can be confident that some of these points are Pareto optimal while others are not. Furthermore we can give sufficient conditions for finding all Pareto optimal allocations.

For the result below we need to consider yet another minor modification of the baseline model and we replace the minimum utility constraint (12) by an equality constraint

$$\sum_{t=0}^{\infty} \delta^t \, [u(c_{2,t}) + v(l_{2,t})] = \underline{U}^2 \tag{28}$$

where $\underline{U}^2$ is restricted so that the set of feasible competitive equilibria satisfying this constraint is non-empty. Let us call this the "modified model number 3", or MM3.

Finally, let us consider yet another modified model, MM4 which consists of solving, for a given $\alpha \in [-\infty, \infty]$

$$\max \sum_{t=0}^{\infty} \delta^t \Big[\; u(c_{1,t}) + v(l_{1,t}) + \alpha \left[u(\lambda c_{1,t}) + v(f(\lambda, l_{1,t})) \; \right] \Big] \tag{29}$$

subject to all competitive equilibrium constraints and the tax limit. Notice that we allow for negative α's and that we consider the case $\alpha = \pm\infty$ as a convention to denote the case where agent 1 or 2 receive no weight.

We show that MM4 can be used to trace a large part of the frontier $\mathcal{F}$ and the Pareto optimal allocations within it.

Given $\alpha \in [-\infty, \infty]$, let $U_i(\alpha)$ be the utility of consumer $i = 1, 2$ in the solution to MM4. Assume

A1 *A solution to MM4 exists for all $\alpha \in [-\infty, \infty]$. Also, $U_i(\alpha)$ is well defined for $i = 1, 2$.*

That a solution exists is satisfied by standard requirements such as that the set of equilibria is non-empty and that the utility functions are bounded. That $U_i(\alpha)$ is well defined amounts to assuming that each α gives a unique utility level for each agent or, equivalently, that $\mathcal{F}$ does not have a linear part.

Proposition 2: *Assume A1.*

1. *given $\alpha \in [-\infty, \infty]$ the solution of MM4 also solves MM3 for $\underline{U}_2 = U_2(\alpha)$.*
2. *given $\alpha \in [-\infty, \infty]$ the solution of MM4 defines a point on the frontier:*

$$(U_1(\alpha), U_2(\alpha)) \in \mathcal{F}$$

3. *if $\alpha \geq 0$ the solution to MM4 is pareto-optimal*
4. *Assume, in addition,*

   A2 *$U_2(\cdot)$ is monotonic decreasing and invertible for $\alpha \in [0, \infty]$.*

   then every Pareto optimal allocation is also the solution of MM4 for some $\alpha \geq 0$

Proof

Fix $\alpha \in [-\infty, \infty]$. To show part 1, let $U_1^{MM3}(\underline{U}_2)$ be the value of the maximum of MM3. By definition, $U_1(\alpha) + \alpha U_2(\alpha)$ is the value of the maximum of MM4. Since the solution to MM3 is feasible in MM4 we have

$$U_1(\alpha) + \alpha U_2(\alpha) \geq U_1^{MM3}(U_2(\alpha)) + \alpha U_2(\alpha)$$

so that $U_1(\alpha) \geq U_1^{MM3}(U_2(\alpha))$. Also, the solution to MM4 is feasible in MM3 for $\underline{U}_2 = U_2(\alpha)$, therefore $U_1(\alpha) \leq U_1^{MM3}(U_2(\alpha))$. This shows that $U_1(\alpha) = U_1^{MM3}(U_2(\alpha))$ or, equivalently, that the maxima of MM4 and MM3 coincide when $\underline{U}_2 = U_2(\alpha)$

Now we prove part 2. For any $\alpha \in [-\infty, \infty]$ we just need to find pairs of utilities that do not belong to the set (27) which are arbitrarily close to $(U_1(\alpha), U_2(\alpha))$. Consider any $\varepsilon > 0$. For $\alpha \geq 0$ the pair of utilities $(U_1(\alpha) + \varepsilon, U_2(\alpha) + \varepsilon)$ is outside the set (27), otherwise it would have been chosen over the optimum in MM4 since it achieves a higher value of the objective. Points such as $(U_1(\alpha) + \varepsilon, U_2(\alpha) + \varepsilon)$ can be made arbitrarily close to $(U_1(\alpha), U_2(\alpha))$ by considering arbitrarily small ε. Therefore $(U_1(\alpha), U_2(\alpha))$ is on the frontier for $\alpha \geq 0$. For $\alpha \leq 0$ a similar argument shows that points such as $(U_1(\alpha) + \varepsilon, U_2(\alpha) - \varepsilon)$ are outside the feasible set and can be made arbitrarily close to $(U_1(\alpha), U_2(\alpha))$.

Part 3: if there is a feasible combination of utilities $(\widetilde{U}_1, \widetilde{U}_2)$ that Pareto dominates $(U_1(\alpha), U_2(\alpha))$ the optimum of MM4 would not be attained at $(U_1(\alpha), U_2(\alpha))$.

For part 4, note that if $(\widehat{U}_1, \widehat{U}_2)$ is a Pareto optimal allocation assumption A2 guarantees that there is an $\widehat{\alpha}$ such that $\widehat{U}_2 = U_2(\widehat{\alpha})$. This and the fact that $(\widehat{U}_1, \widehat{U}_2)$ is Pareto optimal means that $\widehat{U}_1 \geq U_1(\widehat{\alpha})$ so that $\widehat{U}_1 + \widehat{\alpha}\widehat{U}_2 \geq U_1(\widehat{\alpha}) + \widehat{\alpha}U_2(\widehat{\alpha})$. But the fact that the solution to MM4 is attained at $(U_1(\widehat{\alpha}), U_2(\widehat{\alpha}))$ means that the reverse inequality also holds. This means that $\widehat{U}_1 + \widehat{\alpha}\widehat{U}_2 \geq U_1(\widehat{\alpha}) + \widehat{\alpha}U_2(\widehat{\alpha})$ and that the maximum of MM4 for $\widehat{\alpha}$ is attained at $(\widehat{U}_1, \widehat{U}_2)$.

■

Therefore by varying α from plus to minus infinity and maximizing (29) we can trace out points on the frontier of equilibria $\mathcal{F}$ and all points $(U_1(\alpha)), U_2(\alpha))$ for positive α are Pareto optimal. Furthermore, under A2 we are sure that we will find all Pareto optimal allocations in this fashion. The points in $\mathcal{F}$ corresponding to negative α solve MM3 for $\underline{U}^2 = U_2(\overline{\alpha})$ but these equilibria are not Pareto optimal since, as indicated by the negative Lagrange multiplier α, the first agent's utility could be increased in MM3 by also increasing $\underline{U}^2$.

More points on the frontier can be found if the consumers trade places in the objective function (29), that is, if α multiplies the utility of agent 1. Then, by varying α (now the weight of agent 1) from zero to negative infinity again we could find points on the equilibrium frontier that are again *not* Pareto optimal and that are obtained by forcing the planner to give a certain utility to agent one.

There is one caveat: assumptions A1 and A2 need to be checked. Since the feasible set is non-convex the only way to check A1 is to check that there is only one solution to MM4. This can be done numerically by searching for more solutions to the first order conditions as in scores of papers where the maximum is found by searching for all critical points and, if more than one is found, the values of the objective function are compared. We can explore numerically if A2 holds by recorded all utilities for a fine grid of α's and checking that $U_2(\alpha)$ is increasing. In this way we can be confident that we found all Pareto optimal competitive equilibria and that we traced out a "large part" of the frontier $\mathcal{F}$. We did this check for all the examples shown below, the case depicte in Figure 1 is discussed in detail below.

The Pareto optimal-Pareto improving (POPI) plans can be found as those points of $\mathcal{F}$ that have a non-negative α and with utilities that are larger than status quo. Note that non-optimal points on $\mathcal{F}$, i.e. points where $\alpha < 0$, may also be Pareto-improving relative to the status quo. In this case, $\mathcal{F}$ has an increasing part, any POPI plan will strictly improve the utility of one of the agents and it is not possible to shift all the gains to the other agent in a Pareto optimal way. All these concepts will be illustrated in the model we consider in section 4.2, in which the frontier features such a part that Pareto dominates the status quo.

4 Numerical Results

We now present and discuss our numerical results. Details on our computational strategy can be found in appendix B. In the next subsection we discuss how we calibrate the model. We then analyze the case in which labor supply is fixed. With a fixed labor supply the Ramsey equilibrium with a representative agent would set all capital taxes equal to zero, and only labor would be taxed, to achieve the first best without any tax distortion. The tax limits would not be binding in the optimum. As we will see, even with a fixed labor supply, with heterogeneous agents the POPI policy is likely to involve a long transition with high capital taxes. This is because the planner will need to redistribute wealth in favor of the worker in order to ensure that his utility increases relative to the status quo. The planner is willing to lose efficiency and have high capital taxes for many periods in order to achieve a Pareto improving allocation. Since for the homogeneous case there are no distortions whatsoever the fixed labor supply model demonstrates clearly that the need to redistribute in a heterogeneous agent Pareto improving allocation is what drives early capital taxes up. Then we go on to study the optimal policy in the case with a somewhat flexible labor supply. This case is not only studied for generality but because it reveals additional features of optimal plans. In this case, as is well known, even in the homogeneous agent model the planner would like to have high early capital taxes. This compounds with the redistributive effect and it turns out that the planner set high early capital taxes and low labor taxes, for many years after the reform starts. This will induce high labor supply in the early periods, it will achieve faster capital accumulation in the initial periods because the return to capital increases even though capital taxes are still high. We contrast these results with those from an extension of our model in which lump sum transfers are permitted as a way of gaining intuition for the forces at work.

4.1 Calibration

Preference parameters	σ_c	1
	σ_l	3
	B	.76
	δ	.96
Heterogeneity Parameters	ϕ_c/ϕ_w	1.05
	$k_{c,-1}$	5.49
	$k_{w,-1}$	-3.47
Production parameters	α^k	.36
	d	.08
	k_{-1}	1.01
Government spending	g	.13
Tax rates before reform	τ^l	.23
	τ^k	.57
Upper bound on cap. tax rate	$\tilde{\tau}$	.57

Table 1: Parameter Values of the Baseline Economy.

All parameters except for the tax rates remain the same during the policy experiments. We calibrate the model at a yearly frequency. An overview of our parameter choices is provided in table 4.1. Here is a justification of these choices

Preferences: The utility function is as stated in section 2. We choose $\sigma_c = 1$ in keeping with a large part of the literature on capital taxes and in order to use proposition 1. The choice of $\sigma_l = 3$ is for the case of an elastic supply of labor, it avoids labor supply from greatly differing across groups with different wealth, notice that this implies a much lower elasticity of labor supply than many applications of real business cycles.[17] The discount factor δ and B are standard, the latter chosen such that in a corresponding representative agent economy agents would work one third of their time in the steady state before the reform.

Heterogeneity: Our two types of agents are heterogeneous with respect to both their labor efficiency ϕ^j and their initial wealth $k_{j,-1}$. For simplicity we will from now on speak of "workers", indexed w, and "capitalists", indexed c. Capitalists are the group whose ratio of wealth to labor efficiency is higher, i.e. they are rich relative to their earnings potential, but both agents work and save. Note, however, that in absolute terms the capitalists are both richer AND more productive. In the status quo before the reform the heterogeneity parameters of table 4.1 translate into a relative consumption of the workers of $\lambda = 0.4$.

We base our choice of relative labor efficiency and wealth on the analysis of the Panel Study of Income Dynamics performed in Garcia-Milà et al. (1995). They argue that the relevant measure of inequality is the wage/wealth ratio $\phi^j/k_{j,-1}$ and they split their sample in five groups each containing a quintile of the distribution of this ratio across the population. We confine ourselves to only two agents in order to facilitate computations and to better understand the workings of the model, but the policies we find would presumably Pareto improve all agents in the middle categories. The degree of heterogeneity in our calibration approximately corresponds to the two most extreme groups in Garcia-Milà et al. (1995). We think of our two agents as representing the two groups, each with a 20% weight in the population, that would have the most opposing views about how to change labor and capital taxes. It seems reasonable to focus on groups of this size, since a tax reform will be difficult to approve and sustain if it hurts 20% of the population. We

[17] See Garcia-Milà et al. (1995) for a discussion of the tradeoffs in choosing σ_l to match cross-section or time-series evidence.

understate heterogeneity relative to a case where the government would only carry out reforms that improved absolutely all agents, since in the actual population there are some agents in these quintiles with even more different wage/wealth ratios than in our model.

Production: We use a standard yearly calibration for technology. The production function is Cobb-Douglas with a capital income share of $\alpha^k = .36$. There is no productivity growth. The depreciation rate is $d = .08$. Initial aggregate capital is such that the corresponding representative agent economy would be in steady state before the reform.

Government: Before the reform the capital and labor income tax rates are 57% and 23% respectively. These are the average marginal tax rates calculated by McGrattan, Rogerson & Wright (1997) for the period 1947-87. Government spending per period, g, is chosen to balance the budget intertemporally with these tax rates in the status quo. It amounts to about 25% of output. Note that the choice of tax rates in the status quo matters for two reasons. First of all, the capital tax rate influences the status quo steady state (and hence initial) capital stock. Secondly, status quo utilities depend on the tax rates, and thus restrict the scope for Pareto improvements.

We assume that during the reform the capital tax rate can never increase above its initial level so we set $\tilde{\tau}$ equal to status quo capital taxes.

4.2 Pareto Optimal-Pareto Improving Plans: Fixed labor supply

The set of POPI plans deviates from the first best for two reasons. One is the standard reason in models of factor taxation: the need to raise tax revenue discourages the supply of capital and/or labor. The second reason is the lack of non-distortive means of redistribution between the agent types. Since our paper is mostly about the latter, we first analyze a case in which only the redistributive effect is present. We consider a case with a fixed labor supply, where taxes can be chosen so as to not discourage factor supply and, in the homogeneous agent case, the first best without any distortion is achieved. In the next subsection we consider an elastic labor supply.

We assume a fixed labor supply equal to one third. This amounts to taking $\sigma_l \to \infty$ with the scaling parameter B appropriately adjusted so that labor supply stays at one third. Except for σ_l, B, all other parameters are as in table 4.1.

In a model of homogeneous agents capital taxes would be abolished immediately and all revenues would be collected free of distortions from taxes on labor. In a model of heterogenous agents, if the government could stipulate agent-specific lump sum transfers between agents at time zero ($T^w = -T^c$, as introduced at the end of section 2) this would resolve the problem of how to redistribute wealth. Then the first best policy could be achieved for any distribution of welfare gains. But in the case of interest where lump sum redistribution is not possible, deviations from the first best policy with zero capital taxes at all times are necessary for distributive reasons.

In figure 1 we compare the set of POPI plans to the first best in terms of welfare gains.[18] First of all note that assumption A2 is satisfied and that by lowering α the utility of agent 2 goes down, therefore Proposition 2 can be used to trace the whole Pareto optimal points and additional points on the frontier. The line labeled "first best" represents those allocations described in the previous paragraph where agent-specific lump sum transfers are available. Clearly, the absence of transfers significantly reduces the scope for Pareto improvements. All POPI plans depicted in the solid line are inferior to the first best. Why? It turns out that the first best plans that are Pareto improvements upon the status quo would all involve positive transfers to the worker, $T^w > 0$. Absent these transfers, the immediate abolition of capital taxes would severely hurt the worker as has been shown previously in a number of contributions.[19] This is because capital taxes in the status quo are disproportionately borne by the capitalist, and when they are abolished labor taxes have to rise in order for the government to meet its budget constraint. This increase in labor taxes due to of an immediate reform has a strong redistributive effect and - from the perspective of the worker - it would overcompensate the welfare gains arising from increased efficiency. The only thing the planner can do to make the abolition of capital taxes palatable for the worker is to keep capital taxes as high as possible for a long time (the N periods of proposition 1) before setting capital taxes to zero. It turns out that

[18] In all the figures reporting results on welfare, the welfare gains for each agent are measured as the percentage, permanent increase in status quo consumption that would give the agent the same utility as in the optimal tax reform. Therefore, the origin of the graph represents the status quo utility, and the positive orthant contains Pareto improving allocations.

[19] See Correia (1995), Garcia-Milà et al. (1995), Domeij & Heathcote (2004) and Conesa & Krueger (2006).

capital taxes have to be at the upper bound for 10 years (in the POPI plan where the worker gains nothing) to 22 years (in the $\alpha = 0$ POPI plan). Some revenue is then still raised from capital taxes so that labor taxes need not raise all revenue. But this implies a cost in efficiency, because the economy remains distorted for a long time, while it would be non-distorted if lump-sum redistribution were feasible. This is why POPI plans are second best even though taxation would be entirely non-distortive in a homogeneous agent model or if lump sum redistribution were available.[20]

The absence of a lump-sum redistributive instrument not only drives the set of POPI plans away from the first best. It moreover limits the degree to which welfare gains can be shifted towards the worker. This can be seen from the dotted continuation of the POPI line in figure 1. This dotted line is part of the frontier of feasible equilibria $\mathcal{F}$, like the POPI line, however, the allocations along this part of the frontier are not Pareto optimal, and they are found for negative α, corresponding to imposing the constraint with equality (28) as in model MM3 described in section 3.2. The junction of the solid and dotted lines corresponds to the point where $\alpha = 0$. It would be possible to improve both agents by moving from the dotted line to the plan corresponding to $\alpha = 0$, I.e. by forcing the capitalist onto a certain (low) utility level the planner also harms the worker.

It is worthwhile noting that even though the utility loss relative to the first best is small if we only focus on equilibria that leave the worker indifferent and give all the benefits of the reform to the capitalist (i.e., if we focus on points where the frontiers cross the vertical axis of figure 1) the utility loss is very large if we try to give some of the benefits to the worker. The most we can give to the worker is a 2% improvement, which is about one-eighth of the most the worker could gain with lump sum redistribution. There is little to be gained from cutting capital taxes if the worker must enjoy most of the benefits.

Finally, we report the optimal policy under the veil of ignorance, when $\alpha = 1$. This policy gives utilities $(\Pi^k, \Pi^l) = (4.41{,}1.53)$ and, as figure 1 shows it is quite close to the point where $\alpha = 0$ and the capitalist receives zero weight.

[20] Notice that in the case of a fixed labor supply the evolution of labor taxes is undetermined, all that matters is that the net present value of labor taxes balances the government budget constraint given the optimal path for capital taxes found.

4.3 Main model

4.3.1 Welfare frontier and capital tax

We now return to our main model for the rest of the analysis, featuring a somewhat elastic labor supply. The value $\sigma_l = 3$ in table 4.1 amounts to a labor elasticity of about 1/3, still a low value compared to many empirical estimates of about .5 for a representative agent. Figure 2 reports the set of POPI plans in terms of welfare gains. Again, we contrast our main model with the case with optimal agent-specific transfers $T^w = -T^c$. Note, though, that the case with transfers is no longer first best because positive capital and/or labor taxes are needed to raise some tax revenue and pay for the government spending.

Again, the absence of redistributive transfers clearly constitutes an extra constraint on the feasible set and the welfare gains are smaller for Ramsey POPI allocations. But the limits to redistribution are less severe here than with exogenous labor supply. The equilibrium frontier $\mathcal{F}$ (the solid line) is now decreasing in the range of Pareto superior allocations, so that it is now feasible to leave either the worker or the capitalist indifferent relative to status quo without violating Pareto optimality. Not only that, the total loss of welfare relative to the case with transfers is now much lower. If we focus, for example, on points that give equal gain to both agents (the points where each frontier crosses the 45^0 line) we see it is roughly 4% for both agents in the POPI line, only slightly below the 4.25% to be gained by both agents with lump sum redistribution. We conjecture, though, that for sufficiently high σ_l and correspondingly inelastic labor supply the picture would start resembling that of figure 1.

As the distribution of welfare gains varies along the frontier of POPI plans, so do the corresponding capital tax schedule and relative consumption of agents. Qualitatively the properties of capital taxes over time are always the same: As we know from Proposition 1 capital taxes stay at their upper bound for all but the last period of the transition and then they transit to zero with at most one intermittent period. A typical time path for capital taxes is drawn in figure 3. But the length of the transition increases as welfare gains are shifted towards the worker. This is illustrated in the first panel of figure 4 showing the duration of the transition in the vertical axis for each POPI allocation indexed by the welfare gain of the worker on the horizontal axis. We see that the number of periods before capital taxes drop to zero increases from about ten to twenty years as we increase the welfare gain of the worker from zero (i.e., leaving the worker indifferent with status quo) to 9%, which leaves the capitalist indifferent with status quo. Along with the duration of the transition the present value share of capital taxes in government revenues increases from 15% to 24%, as the second panel in figure 4 reveals.[21] This is the clue to why a longer period of high capital taxes is beneficial for the worker: The worker contributes to the public coffers primarily through labor taxes, which means his burden in the long run stands to increase through the reform, while the capitalist's long run burden decreases. The earlier capital taxes are suppressed, the more revenue has to be raised from labor taxes and the bigger the relative tax burden of the worker.

The final panel in figure 4 depicts α, the multiplier on the minimum utility constraint for the worker, and λ, the ratio of the worker's consumption to the capitalist's in equilibrium. We put these two graphs in the same picture because in the standard case in dynamic models, under log utility, without distorting taxes and with complete markets, we would have $\alpha = \lambda$. More precisely, this equality holds in a first best situation, without distorting taxation and no distributive conflict ($\Delta_1 = \Delta_2 = 0$), if the upper bound on capital taxes never binds ($\gamma_t = 0 \; \forall t$). In our second best world, by contrast, as we increase the welfare of the worker the marginal cost of doing so (as measured by α) explodes, while his consumption share increases only mildly. In fact, it always remains very close to its value in the status quo, which is 0.4. This shows that it is very difficult to alter the ratio of consumption (λ) even if the planner cares very differently about the agents given that it has access only to proportional taxes and agent-specific lump sum taxes are not available.

If optimal lump sum transfers were possible, the graphs in figure 4 would look very different. We have computed that for all Pareto optimal allocations capital taxes would be suppressed after 5-6 years, and the share of capital taxes would always be 0.11. The multiplier on the worker's utility constraint α would increase very little with $\Pi(worker)$, while λ would rise much more than without the transfer. This is because in this case the redistribution can be achieved with agent-specific lump sum taxes independently of the fact that the planner lowers quickly capital taxes to achieve aggregate efficiency. The

[21] For comparison, the share of capital taxes in revenues is about .43 in the status quo.

policies and the path of the economy would hardly depend on the distribution of the gains from reform. That shifting welfare gains and consumption between agents would be much easier, as indicated by the behavior of α and λ.[22]

Focusing on Pareto improving allocations means that the unit of interest is the utility that each agent achieves through various tax reforms. Under this view, the weight α is just a Lagrange multiplier determined in equilibrium, and it measures the cost of enforcing the minimum utility constraint. The fact that α has to increase so much to achieve a small redistribution is just a reflection of the difficulties that the planner finds in redistributing wealth from one agent to the other when only capital or labor taxes are available.

Another way of looking at α is as the relative weight that the worker receives in the welfare function of the government. This suggests to interpret α as a measure of the bias of the social planner in favor of the workers. In particular, if one were to focus on the optimal allocation under the "veil of ignorance", since both types of agents are equally abundant in the economy, the relevant policy would be the one corresponding to $\alpha = 1$. This also corresponds to a model of probabilistic voting where both agents are equally influential. Many recent papers with on dynamic optimal policy with heterogeneous agents use this welfare function. As can be seen from the bottom panel in Figure 4 the optimal policy for $\alpha = 1$ corresponds to a welfare gain by the worker of about 8%. This point has also been depicted in the frontier of equilibria in figure 1. By chance, the optimal policy under the veil of ignorance happens to be Pareto improving. It gives most of the benefit of the reform to the worker. For this welfare function capital taxes are zero after 18 years.

In appendix D we show that the main features described here are robust to changes in the parameter values when some of the parameters are changed in one direction at a time from the benchmark case.

4.3.2 The time path of the economy

The evolution of capital, labor supply, the labor tax rate, and the government deficit are pictured in figure 5. First, note that qualitatively the paths are very similar across the set of POPI plans. The horizontal shifts in the graphs occur because the more a plan

[22] Note that even with redistributive lump sum taxes we do not obtain $\alpha = \lambda$. This only holds in optimal allocations when there is no distortionary taxation.

benefits the worker the longer capital taxes remain at their initial level. The kinks in the paths of labor taxes and government deficit occur precisely in the intermediate period when capital taxes transit from their maximum to zero.

The most surprising observation is, perhaps, that labor taxes should be initially lowered, and they should remain low for a long time. The reason for this behavior is the following: the planner wants to frontload capital taxes for the usual reason described at length in the literature that early capital taxes imply taxing capital that is already in place.[23] Therefore, it is optimal to keep capital taxes at the upper limit in the first few periods and then let them go to zero. With such high capital taxes investors would not invest much. But the government has another instrument that can be used to boost output and capital accumulation in the early periods. The government can lower labor taxes, inducing an increase in the labor supply, causing the return of capital to go up, increasing investment in the initial periods, and achieving a faster convergence to the optimal long run capital/labor ratio compatible with zero capital taxes. The upper right panel in figure 5 shows that aggregate labor supply is very high in the early periods.[24] Note that the accumulation of capital accelerates around the period that capital taxes become zero, as can be seen by comparing the kink in the graph for labor taxes with the capital accumulation graph. Therefore, eventually the zero capital tax is the one promoting growth and helping the economy converge to the new steady state. Absent this backloading of labor taxes capital would initially grow only to the extent that the expectation of low capital taxes in the distant future raises incentives to save early on. In this case capital accumulation would be much slower, as in the fixed labor supply case of section 4.2. Therefore, the low early labor taxes are an instrument to induce investment in the early periods that can be used in the case of an elastic labor supply.

A similar result of low early labor taxes has been found in models of homogeneous agents.[25] The same pattern can be observed in our model if optimal transfers are allowed, we have computed that in the case of agent-specific lump sum transfers $T^w = -T^c$ the

[23] For example, Jones, Manuelli and Rossi (1993) describe in detail this issue in several models with homogeneous agents.

[24] Capitalists, who also have higher labor productivity, always work less than workers.

[25] For example, section III of Jones, Manuelli and Rossi (1993) shows a model where labor taxes should be very negative and capital taxes should be very high in the first period only.

period of low labor taxes would be much shorter, of about 5-6 years, matching the lower duration of the transition to zero capital taxes. But implementing this policy without the lump sum transfers would leave the workers worse off relative to status quo. Redistributive concerns lengthen the transition three or four times, as described in the previous paragraph.

It is interesting that the redistributive effect and the effect of promoting growth go in the same direction: they both induce the planner to set low initial labor taxes. This explains why with flexible labor supply the POPI frontier is closer to the frontier with optimal transfers, as shown in Figure 2, than it was in the fixed labor supply case. With an elastic labor supply the desire to boost investment early on is not in conflict with the redistribution objective.

A somewhat surprising pattern that emerges from the pictures is that the long run labor tax rate is higher for the policy that favors the worker. This may seem paradoxical because the worker is interested in low labor taxes. Note though, that even though the *long run* labor tax rate is higher if the worker is favored, the initial cut is even stronger for these policies, so that the share of labor taxes in the total *present value* of revenues is lower for policies that favor more the worker, as figure 4 showed. This suggests that the long run labor tax rate is high for two reasons. First, when capital taxation is abandoned late the initial boost to capital accumulation comes mainly from extremely low initial labor taxes. I.e. the backloading of labor taxes is strongest in these cases. Second, the long run labor supply is lower the later capital taxes are suppressed, while the gross wage is always the same.[26]

Since government expenditures are constant, the low initial labor taxes translate into government deficits. Only as labor taxes rise and output grows the government budget turns into surplus. Once capital taxes are suppressed and revenues fall again, the government deficit quickly reaches its long run value which can be positive or negative depending on whether during the transition the government accumulated wealth or not. We can see from figure 5 that most POPI policies imply that the government runs a primary surplus

[26]Since the long run real return on capital is determined by the rates of time preference and depreciation and the production function is Cobb-Douglas, the long run capital-labor ratio and wage are independent of the policy - as long as capital taxes are zero eventually.

in the long run. This implies that the government is in debt in the long run, because the primary surplus is needed to pay the interest on debt. Therefore, for most POPI tax reforms low taxes in the initial periods generate a positive level of long run government debt.

This feature of the model is quite different from that of Chamley, where the government accumulated savings in the early periods to lower the labor tax bill in the long run. Here, the early drop in labor taxes is financed in part with long run government debt, showing that one possible reason for government debt is to finance the initial stages of a reform.

4.4 Related issues

Our main intention has been to comment on the Chamley/Judd result. Therefore, we have stayed as close as possible to their model. Now we explore several variations of the model to consider issues of progressive taxation, political sustainability of equilibrium, and time consistency.

4.4.1 Progressive taxes

Given that we set out to analyze the consequences of distributive concerns for optimal tax policy, it might strike the reader as very restrictive to allow proportional factor taxation only. After all, one of the prime instruments of redistribution in the real world is progressive taxation so it is natural to ask if allowing for a progressive tax code would solve the redistributive concerns and cause the economy to be closer to the first best. We therefore now consider an extension of our model that allows for non-proportional taxes in a simple way.

We assume that the planner can choose a uniform lump-sum payment $\mathcal{D}$ that is paid in period zero uniformly across all agents. Following Werning (2007), under complete markets this is equivalent to a fixed deductible from the tax base in each period. A positive $\mathcal{D}$ means progressive taxation. Introducing this in the model implies that we need to add $u'(c_{1,0})\left[\Delta_1 + \Delta_2\right]\mathcal{D}$ to the $\mathbf{W}$-term in equation (16). We then let the planner maximize over $\mathcal{D}$ additionally.

We find that if we restrict our attention to non-negative $\mathcal{D}$ (progressive taxation), the optimal choice is to set $\mathcal{D}$ =0. Therefore, the government will choose *not* to use this progressive instrument.

The reason for this result is the following: there are two forces at work in the determination of the optimal $\mathcal{D}$. On the one hand, redistributive concerns would advise the government to choose a positive $\mathcal{D}$, since capitalists are richer. But a negative $\mathcal{D}$ is equivalent to a lump-sum tax, and it allows to raise revenue in a distortion-free manner. In the standard case of a representative agent model, where only this second force is present, the first best can be achieved by choose a negative $\mathcal{D}$ big enough (in absolute value) to raise all government revenue ever needed. In our model with heterogeneous agents it turns out that the second force is stronger.

We find that, if it could, the government would set a negative $\mathcal{D}$. How can this be Pareto improving? The reason is that the government now redistributes by choosing very negative labor taxes for many periods. In fact, the present value of revenues from labor taxes is not only negative but even bigger in absolute value than the revenue from capital taxes. The transition is faster than in the main model, between 4 and 14 periods in both extremes of the POPI frontier. Welfare gains are larger than in the case with optimal transfers.

This optimal tax scheme (negative $\mathcal{D}$ and negative labor taxes) is Pareto improving only because we did not consider agents of different wealth within each type of agent in our calibration. In the real world some agents have a high wage/wealth ratio who are rich (say, some young stockbrokers) and agents with a low wage/wealth ratio who are poor (say, some farmers in economically depressed areas). We calibrated according to wage/wealth ratios because following Garcia-Milà et al. (2000) this is appropriate when only proportional taxes are allowed, but once progressive taxation is considered the total income of the agent is also relevant. Therefore, a careful study of progressive taxation should introduce total income in the calibration, in that case the optimal scheme described above would be unlikely Pareto improving. This is left for future research. But the results in this subsection show that progressive taxation will have many difficulties in solving the redistribution problem.

4.4.2 The evolution of wealth and welfare and time consistency

One might conjecture that the welfare of workers and capitalists drifts apart in the long run, with capitalists profiting from the abolition of capital taxes and workers suffering from high labor taxes in the long run. It might seem that such a scenario would render the tax reform politically unsustainable. We now study this issue, first informally by exploring the evolution of welfare and wealth and then more formally by addressing issues of time consistency.

The time paths of agents' wealth and welfare are plotted in figure 6. Welfare increases along with the accumulation of capital, and - contrary to the conjecture - both agents' welfare evolves more or less in lock step. The reason is that, by the competitive equilibrium conditions (7) both their relative consumption and their relative leisure are roughly constant over time. Therefore it is not the case that workers will lose dramatically when capital taxes finally drop to zero.

This is an implication of the permanent income hypothesis. Agents' income net of taxes varies through time, so agents will save or dissave in order to smooth consumption and hours. The smooth time path of welfare is made possible by a less smooth path of individual wealth. Since the worker's main contribution to public coffers is due in later periods when labor taxes are high and in the early years of the new policy he benefits from extremely low labor taxes, he accumulates wealth to provide for the higher tax burden he will face later on. The capitalist's tax burden, by contrast, tends to decrease over time since initial capital taxes are very high and they are later suppressed. By deferring wealth accumulation until his tax burden drops, he can afford a smoother consumption profile.

The fact that welfare of both agent types increases over time in a similar fashion suggests that the solution is, in some sense, politically sustainable. We can study if the solution we found is time consistent more formally.[27] An issue to be addressed in this case is what is the objective function of the planner at the time of reoptimization at some future date. One alternative is to assume that the planner uses the same welfare function implied by the initial Lagrange multiplier on (15) α. Another alternative is to assume that the weight α switches through time and it is set according to a bargaining scheme

[27] The literature on time consistency in models with heterogenous agents is not very large. One exception is Armenter (2004).

or political competition. Another option is to assume that consensus is required so that a Pareto improvement is needed for reoptimization to take place.

We perform some numerical checks that give a partial answer to the question of time consistency. We assume that the optimal plan is followed for M periods and then the planner reoptimizes taking $k_{g,M-1}, k_{w,M-1}, k_{c,M-1}$, and τ_M^k as given. We then check whether the reoptimized solution differs from the remaining path under the original solution for various fixed weights and when consensus is required.[28]

Let us first consider constant welfare weights. This approach could be motivated with reference to a political economy set-up with one-off probabilistic voting where the political power of each agent is constant through time. Reoptimization then means that voters unexpectedly get the chance to vote again about the best capital/labor tax policy in period M.[29] Under probabilistic voting a weighted sum of utilities is maximized, with the weights determined by population weights and the degree to which a voter group votes on grounds of policy rather than ideology.[30] For example, when both of our equal sized groups are equally easy to sway, they each receive the same welfare weight ($\alpha = 1$), and the situation is the same as under the "veil of ignorance". Under this approach, our optimal plans are not time consistent. In all our calculations reoptimization with constant welfare weights leads to an extension of the period during which capital taxes remain high. This raises the utility of the worker relative to his continuation utility under the original policy and reduces that of the capitalist. For example, for the case $\alpha = 1$, if we reoptimize in period $M = 5$, the total duration of the transition (counting from $t = 0$) increases from 18 to 23 years. Relative to the continuation utility under the original optimal plan the worker experiences a welfare gain of 3%, while the capitalist loses 2.2%. Therefore, in this sense, the worker would prefer to delay yet again the elimination of capital taxes.

We also consider the case where the reoptimization takes place only if a Pareto improving allocation can be found at time M relative to the agents' continuation utilities at the period of reoptimization. This would correspond to a situation where the tax reform

[28] Here, unlike in period zero, government may have issued bonds during the first M periods so $k_{g,M-1}$ may be non-zero.

[29] If agents expected to vote repeatedly, we would have to consider a dynamic voting model. This would complicate the model and it would mean going further away from the original setup of Chamley.

[30] For a brief introduction to probabilistic voting see Persson & Tabellini (2000), ch. 3.

announced in period 0 can be changed *only* if there is consensus across all agents to do so.

From our numerical experiments it seems impossible to make one agent strictly better off without hurting the other, so that reoptimizing with consensus always leads to the confirmation of the original plan in terms of taxes and allocations. Only the time-invariant multipliers α and Δ_i change. Notice that this is compatible with the results of constant weights discussed in the previous paragraph, since in the reoptimized policy with constant weights involves a utility loss of the capitalist and it is, therefore, not Pareto improving in period M. The time-variant multipliers μ_t and γ_t are rescaled by a factor $\frac{1+\tilde{\alpha}}{1+\alpha}$, where the tilde indicates the reoptimized solution. Moreover, we have the relationship $\gamma_{M-1} = \frac{1+\alpha}{1+\tilde{\alpha}}\left(\tilde{\Delta}_1 k_{1,M-1} + \tilde{\Delta}_2 k_{2,M-1}\right)$. Inspection of the first order conditions shows that the remainder of the original optimal plan satisfies the first order conditions of the reoptimization problem if these relationships between the multipliers hold and $\tilde{\alpha}$ and the Δ_i are appropriately chosen. Interestingly, $\tilde{\alpha}$ always turns out smaller than α. For instance, in the above mentioned case of $\alpha = 1$ and reoptimization in $M = 5$ the continuation utilities are respected if $\alpha = 0.603$. Thus, effectively the influence of the worker on the solution under consensus reform has to be lower at the point of reoptimization for the original solution to be time consistent.

This suggests that in order to sustain the tax reform it is not necessary to write the reform as part of a constitution that cannot be changed forever at any cost. It is enough to require that the constitution can only be changed under wide consensus for the tax reform to be sustainable.[31]

5 Conclusion

We find that there is, most definitely, an equity/efficiency trade-off in the determination of capital and labor taxes. Capital taxes should be zero in the long run, but this is an optimal Pareto improving policy only if capital taxes are very high, and labor taxes very low, for a very long time after the reform starts. The government typically accumulates debt in the long run in order to finance the initial cut in labor taxes. Lower initial labor

[31] This result is reminiscent of the one found by Armenter (2004) in a different model.

taxes are necessary for two reasons: first, to redistribute wealth in favor of the worker so as to ensure that he/she also gains from the reform and, second, to boost investment in the initial periods. These features of the optimal policy remain in the special case when the planner has a welfare function that weighs all agents equally as in the optimal policy under the veil of ignorance.

Many of our results are numerical, for a given calibration of heterogeneity according to wage/wealth ratios. The results are robust to many variations in parameter values and even to the introduction of progressive taxation. If labor supply is inelastic it is very costly to make the worker enjoy significant benefits from the capital tax cut. The solution seems to be time consistent if consensus is required at the time of reoptimization, suggesting that the tax reform is credible if it can only be overturned when all agents agree. On the other hand, the solution is time inconsistent if reoptimization takes place with a welfare function that has constant weights.

While the Chamley/Judd result may have discouraged some economists from studying optimal capital and labor taxes with heterogeneous agents, we find that issues of redistribution are crucial in designing optimal policies involving capital/labor taxes. Therefore, research on these issues should be encouraged, both from an empirical and theoretical point of view.

One avenue for research is to study other policy instruments that may be used to compensate the workers for the elimination of capital taxes. For example, certain types of government spending or other tax cuts could have this role. More empirical work on the relevant aspects of heterogeneity so that issues of progressivity can be addressed carefully is certainly needed. The transition in our model is very long. Less-than-full credibility and less-than-fully rational expectations might render this policy not very effective in practice. Introducing issues of partial credibility, time consistency, learning about expectations and political economy would therefore be of interest and might influence the picture on what an optimal policy should do.

A general methodological lesson is that one needs to go beyond long run analysis in order to make any policy recommendations. A researcher looking only at the long run in our model would give the wrong recommendation that capital taxes should be suppressed to achieve a Pareto improvement, while this is not at all what should be done for many periods. In particular, the nowadays fashionable 'timeless perspective', which is only interpretable as the behavior of optimal policy in the steady state, would give exactly this wrong recommendation. The steady state analysis usually found in the literature is, therefore, only the beginning of a normative analysis, but studying the transition is not to be overlooked.

References

Aiyagari, S. (1995), 'Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting', *Journal of Political Economy* **103**(6), 1158–75.

Armenter, R. (2004), Redistribution, time consistent fiscal policy and representative democracy. Working Paper, New York Fed.

Atkeson, A., Chari, V. & Kehoe, P. (1999), 'Taxing capital income: A bad idea', *Federal Reserve Bank of Minneapolis Quarterly Review* **23**(3), 3–17.

Bassetto, M. (1999), Optimal fiscal policy with heterogeneous agents. Working Paper, Chicago Fed.

Chamley, C. (1986), 'Optimal taxation of capital income in general equilibrium with infinite lives', *Econometrica* **54**, 607–622.

Chamley, C. (2001), 'Capital income taxation, wealth distribution and borrowing constraints', *Journal of Public Economics* **79**(1), 55–69.

Chari, V., Christiano, L. & Kehoe, P. (1994), 'Optimal fiscal policy in a business cycle model', *Journal of Political Economy* **102**(4), 617–52.

Chari, V. & Kehoe, P. (1999), Optimal fiscal and monetary policy, *in* J. Taylor & M. Woodford, eds, 'Handbook of Macroeconomics', Vol. 1, Elsevier, chapter 26, pp. 1671–1745.

Conesa, J. & Krueger, D. (2006), 'On the optimal progressivity of the income tax code', *Journal of Monetary Economics* **53**(7), 1425–1450.

Correia, M. I. H. (1995), Efficiency and equity: Is there a trade-off?, CEPR Discussion Papers 1218, C.E.P.R. Discussion Papers.

Domeij, D. & Heathcote, J. (2004), 'On the distributional effects of reducing capital taxes', *International Economic Review* **45**(2), 523–554.

Faraglia, E., Marcet, A. & Scott, A. (2006), In search of a theory of debt management. working paper.

Flodén, M. (2006), Why are capital income taxes so high?, Working Paper Series in Economics and Finance 623, Stockholm School of Economics.

Garcia-Milà, T., Marcet, A. & Ventura, E. (1995), Supply side interventions and redistribution, Economics Working Papers 115, Department of Economics and Business, Universitat Pompeu Fabra.

Giannitsarou, C. (2006), 'Supply-side reforms and learning dynamics', *Journal of Monetary Economics* **53**(2), 291–309.

Jones, L., Manuelli, R. & Rossi, P. (1993), 'Optimal taxation in models of endogenous growth', *Journal of Political Economy* **101**(3), 485–517.

Judd, K. (1985), 'Redistributive taxation in a simple perfect foresight model', *Journal of Public Economics* **28**, 59–83.

Lau, L., Qian, Y. & Roland, G. (2001), 'Reform without losers: An interpretation of china's dual-track approach to transition', *Journal of Political Economy* **108**(1), 120–143.

Lucas, R. (1990), 'Supply side economics: An analytical review', *Oxford Economic Papers* **42**, 293–316.

Marcet, A. & Marimon, R. (1998), Recursive contracts, Technical report.

Marcet, A., Obiols-Homs, F. & Weil, P. (2007), 'Incomplete markets, labor supply and capital accumulation', *Journal of Monetary Economics* **54**, 2621–2635.

McGrattan, E., Rogerson, R. & Wright, R. (1997), 'An equilibrium model of the business cycle with household production and fiscal policy', *International Economic Review* **38**(2), 267–90.

Niepelt, D. (2002), Tax smoothing versus tax shifting. Working Paper, IIES Stockholm.

Persson, T. & Tabellini, G. (2000), *Political Economics. Explaining Economic Policy*, MIT Press.

Werning, I. (2007), 'Optimal fiscal policy with redistribution', *Quarterly Journal of Economics* **122**(2), 925–967.

A The maximization problem and first order conditions

Using the derivations in section 2, the maximization problem to be solved becomes

$$\max_{\lambda,\{c_t^1,k_t,l_t^1\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \delta^t \, [u(c_{1,t}) + v(l_{1,t})] \tag{30}$$

$$\text{s.t.} \qquad u'(c_{1,t}) \geq \delta \; u'(c_{1,t+1})(1 + (r_{t+1} - d)(1 - \tilde{\tau})) \qquad \text{for all } t \tag{31}$$

$$\frac{1+\lambda}{2} c_{1,t} + g + k_t - (1-d)k_{t-1} = F\left(k_{t-1}, \frac{l_{1,t} + f(\lambda, l_{1,t})}{2}\right) \qquad \text{for all } t \tag{32}$$

$$\sum_{t=0}^{\infty} \delta^t \, [u(\lambda c_{1,t}) + v(f(\lambda, l_{1,t}))] \geq \underline{U}^2 \tag{33}$$

$$\sum_{t=0}^{\infty} \delta^t \left(u'(c_{1,t})c_{1,t} + v'(l_{1,t}) \; l_{1,t}\right) = u'(c_{1,0}) \; k_{1,-1}(1 + (r_0 - d)(1 - \tau_0^k)) \tag{34}$$

$$\sum_{t=0}^{\infty} \delta^t \left(u'(c_{1,t})\lambda c_{1,t} + \frac{\phi_2}{\phi_1} v'(l_{1,t}) \; f(\lambda, l_{1,t})\right) = u'(c_{1,0}) \; k_{2,-1}(1 + (r_0 - d)(1 - \tau_0^k)) \tag{35}$$

letting $\alpha, \Delta_1, \Delta_2$ be the Lagrange multipliers for the constraints involving discounted sums (33), (34) and (35), the Lagrangian is given by (16).

The first order conditions for the Lagrangian are:

- for consumption, $t > 0$:

$$u'(c_{1,t}) + \alpha\lambda u'(\lambda c_{1,t}) + (\Delta_1 + \lambda\Delta_2)[u'(c_{1,t}) + u''(c_{1,t})c_{1,t}] +$$
$$\gamma_t u''(c_{1,t}) - \gamma_{t-1} u''(c_{1,t})(1 + (r_t - d)(1 - \tilde{\tau})) = \mu_t \frac{1}{2}(1 + \lambda)$$

- for consumption, $t = 0$:

$$u'(c_{1,0}) + \alpha\lambda u'(\lambda c_{1,0}) + (\Delta_1 + \lambda\Delta_2)[u'(c_{1,0}) + u''(c_{1,0})c_{1,0}] - u''(c_{1,0})\big((\Delta_1 k_{1,-1} + \Delta_2 k_{2,-1})(1 + (r_0 - d)(1 - \tau_0^k)) + \Delta_1 m_{1,,1} + \Delta_2 m_{2,-1}\big) + \gamma_0 u''(c_{1,0}) = \mu_0 \frac{1+\lambda}{2}$$

- for labor, $t > 0$:

$$v'(l_{1,t}) + \alpha v'(f(\lambda, l_{1,t}))f'(\lambda, l_{1,t}) + \Delta_1[v'(l_{1,t}) + v''(l_{1,t})l_{1,t}] + \Delta_2 \frac{\phi_2}{\phi_1}[v'(l_{1,t})f'(\lambda, l_{1,t}) + v''(l_{1,t})f(\lambda, l_{1,t})] - \gamma_{t-1}u'(c_{1,t})(1 - \tilde{\tau})F_{k,e}(k_{t-1}, e_t)\frac{1}{2}(\phi_1 + \phi_2 f'(\lambda, l_{1,t})) = -F_e(k_{t-1}, e_t)\frac{1}{2}(\phi_1 + \phi_2 f'(\lambda, l_{1,t}))\mu_t$$

- for labor, $t = 0$:

$$v'(l_{1,0}) + \alpha v'(f(\lambda, l_{1,0}))f'(\lambda, l_{1,0}) + \Delta_1[v'(l_{1,0}) + v''(l_{1,0})l_{1,0}] + \Delta_2 \frac{\phi_2}{\phi_1}[v'(l_{1,0})f'(\lambda, l_{1,0}) + v''(l_{1,0})f(\lambda, l_{1,0})] - (\Delta_1 k_{1,-1} + \Delta_2 k_{2,-1})u'(c_{1,0})F_{k,e}(k_{t-1}, e_t)\frac{1}{2}(\phi_1 + \phi_2 f'(\lambda, l_{1,t}))(1 - \tau_0^k) = -F_e(k_{t-1}, e_t)\frac{1}{2}(\phi_1 + \phi_2 f'(\lambda, l_{1,t}))\mu_0$$

- for capital, $t \geq 0$:

$$\mu_t + \gamma_t \delta u'(c_{1,t+1})(1 - \tilde{\tau})F_{k,k}(k_t, e_{t+1}) = \delta\mu_{t+1}(1 + F_k(k_t, e_{t+1}) - d)$$

- for λ:

$$\sum_{t=0}^{\infty} \delta^t \Big[\alpha\ (u'(\lambda c_{1,t})c_{1,t} + v'(f(\lambda, l_{1,t}))f_\lambda(\lambda, l_{1,t})) + \Delta_2\big(u'(c_{1,t})\lambda c_{1,t} + \frac{\phi_2}{\phi_1}v'(l_{1,t})f_\lambda(\lambda, l_{1,t})\big) - \mu_t \frac{1}{2}\big(c_{1,t} - F_e(k_{t-1}, e_t)\phi_2 f_\lambda(\lambda, l_{1,t})\big)\Big] = 0$$

B Computational strategy: Approximation of the time path

1. Fix T as the number of periods after which the steady-state is assumed to have been reached. (We use $T = 150$.)
2. Propose a $3{*}T{+}3$-dimensional vector $X = \{k_0, ..., k_{T-1}, l_0, ..., l_{T-1}, \gamma_0, ..., \gamma_{T-1}, \Delta_1, \Delta_2, \lambda\}$. (This is not the minimal number of variables to be solved for as a fixed point problems. $2{*}T{+}3$ would be sufficient, however, convergence is better if the approximation errors are spread over a larger number of variables.)
3. With k_{-1} and g known, find $\{c_t, F_{kt}, F_{lt}, F_{klt}, F_{kkt}\}$ from the resource constraint and the production function.
4. Calculate $\{\mu_t\}$ from the FOC for labor.
5. Calculate $\{\gamma_t\}$ from the FOC for consumption, making use of $\{\mu_t\}$ and the guess for $\{\gamma_t\}$ from the X-vector. (The guess is plugged into γ_{t-1}, γ_t is backed out.)
6. Form the $3 * T + 3$ residual equations to be set to 0:
   - The FOC for capital (Euler equation) has to be satisfied. (T equations)
   - The vector $\{\gamma_t\}$ has to converge, i.e. old and new guess have to be equal. (T equations)
   - Check for each period whether the constraint on τ^k is satisfied. If yes, impose $\gamma_t = 0$. Otherwise, the constraint on capital taxes has to be satisfied with equality. (T equations)
   - The remaining 3 equations come from the present value budget constraints (PVBC) and the FOC for λ. The discounted sums in the PVBCs are calculated using the time path of the variables for the first T periods and adding the net present value of staying in steady-state forever thereafter.
7. Iterate on X to set the residuals to 0. (We use Broyden's algorithm to solve this $3 * T + 3$-dimensional fixed point problem.)

C Recursive Formulation

Formally, the structure of Marcet & Marimon (1998) does not apply starting at period $t = 0$ because the terms grouped in $\mathbf{W}$ in the above Lagrangian have some endogenous variables (labor and consumption) that appear differently in period $t = 0$ than in all remaining periods. The model is recursive only after period 0.

Given constants Δ, λ, α it can be shown that a time invariant policy function F

$$\begin{bmatrix} c_t^1 \\ k_t \\ l_t^1 \\ \gamma_t \end{bmatrix} = F(k_{t-1}, \gamma_{t-1}) \qquad t \geq 1 \tag{36}$$

$$= F_0(k_{-1}) \qquad\qquad t = 0 \tag{37}$$

gives the solution for any initial value of k_{-1} and for $\gamma_0 = 0$. That is, the solution is recursive after period 1 if the multipliers γ are introduced as co-state variables. The dependence of F on the Δ's, α and λ is left implicit here.

To show (36), fix the constants given Δ's, λ and α, first solve the problem from period $t = 1$ onwards

$$\begin{aligned} \max_{\{c_t^1, k_t, l_t^1\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \delta^t \Big[& u(c_{1,t}) + v(l_{1,t}) + \alpha(u(\lambda c_{1,t}) + v(f(\lambda, l_{1,t}))) + \\ & \Delta_1(\; u'(c_{1,t})c_{1,t} + v'(l_{1,t}) \; l_{1,t} \;) + \\ & \Delta_2 \; \left(u'(c_{1,t})\lambda c_{1,t} + \frac{\phi_2}{\phi_1} v'(l_{1,t}) \; l_{2,t} \right) + \\ & \gamma_0 u'(c_{1,1})(1 + (r_1 - d)(1 - \tilde{\tau})) \Big] \end{aligned}$$

subject to the tax limit for all periods and feasibility, and for fixed values of Δ's, α, λ, γ_0 and k_0. Notice that in this problem the series to be found starts at $t = 1$, and the choices in period zero are taken as given. Given the optimal choices in period zero for k, γ, we now consider the maximization for the periods $t > 0$. Solution of this model is a special case of the framework in Marcet and Marimon.

Let $W(\gamma_0, k_0)$ be the value at the optimum of the objective function of the above maximization problem. The dependence on the constants Δ's, λ and α is left implicit. Now we solve for the period zero quantities by finding

$$\begin{aligned}\max_{(c_0^1,k_0,l_0^1)} u(c_{1,0}) + v(l_{1,0}) + \alpha[u(\lambda c_{1,0}) + v(f(\lambda, l_{1,0}))] + \\ \Delta_1[\ u'(c_{1,0})c_{1,0} + v'(l_{1,0})\ l_{1,0}\ ] + \\ \Delta_2\ \left(u'(c_{1,0})\lambda c_{1,0} + \frac{\phi_2}{\phi_1} v'(l_{1,0})\ f(\lambda, l_{1,0})\right) + \\ \delta W(\gamma_0, k_0)\end{aligned}$$

subject to tax limit and feasibility in period zero. It is easy to see that the solution of this maximizes the original problem by combining the arguments in Chari, Christiano & Kehoe (1994) and the ones in Marcet & Marimon (1998). Solving this last model gives rise to the decision function in period zero (37).

D Sensitivity analysis

To check the sensitivity of our results to the choice of parameters we separately vary the preference parameters, tax rates in the status quo, and heterogeneity in each case leaving the remaining parameters as in the main calibration. We always find the same qualitative properties of the optimal policy that we described in section 4. Table D summarizes the results by reporting the duration of the transition and the revenue share of capital taxes for the extreme points of the set of POPI plans. The general pattern is that higher risk aversion, makes the transition longer with a correspondingly higher share of capital taxes in revenues. Higher initial (and maximum) capital taxes and less heterogeneity (i.e. high λ_{SQ}) make the policy more sensitive to the distribution of welfare gains in the sense that the difference in duration and revenue share between the two extreme ways of distributing gains becomes larger.

Calibration	$\Pi(capitalist) = 0$ duration	rev. share of τ^k	$\Pi(worker) = 0$ duration	rev. share of τ^k
benchmark	20	24%	10	15%
$\gamma_c = 0.5$	16	21%	9	9%
$\gamma_c = 2$*	18	22%	12	18%
$\tau^k_{SQ} = 0.4$	18	16%	12	12%
$\tau^k_{SQ} = 0.65$	21	28%	8	15%
$\tau^l_{SQ} = 0.18$	20	27%	10	17%
$\tau^l_{SQ} = 0.28$	20	21%	10	13%
$\lambda_{SQ} = 0.31$**	19	24%	10	16%
$\lambda_{SQ} = 0.5$	21	25%	8	13%

The column entitled 'Calibration' indicates which parameter has been reset to which value. All other parameters are as in the benchmark case. The subscript 'SQ' refers to the status quo before the reform.

* In this case, like with an inelastic labor supply, the policy that has the maximum utility gain for the worker still leaves the capitalist with a strict welfare gain. We therefore report this policy rather than the one that would have $\Pi(capitalist) = 0$.

** Here heterogeneity parameters correspond to the two most extreme groups in GMV.

Table 2: Sensitivity analysis

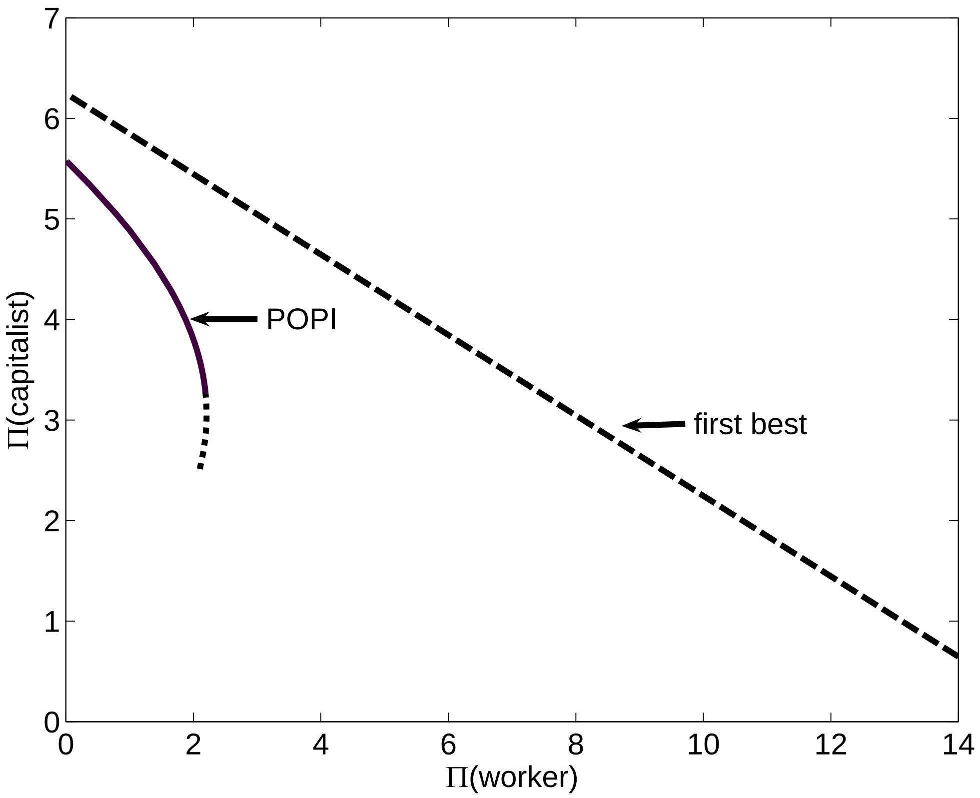


Figure 1: The frontier of Pareto improving feasible equilibria in terms of welfare gains (in %) in the model with fixed labor supply.

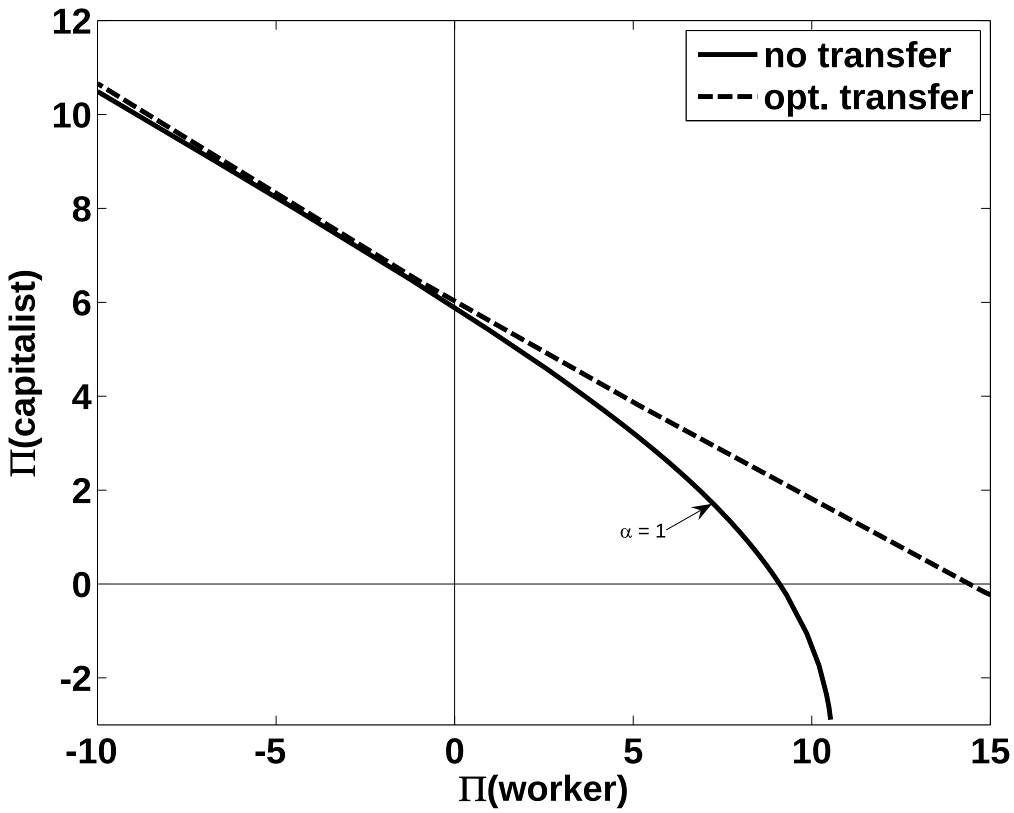


Figure 2: The frontier of feasible equilibria in terms of welfare gains (in %) in the baseline model and with transfers
The point $\alpha = 1$ corresponds to the policy under the so called 'veil of ignorance'.

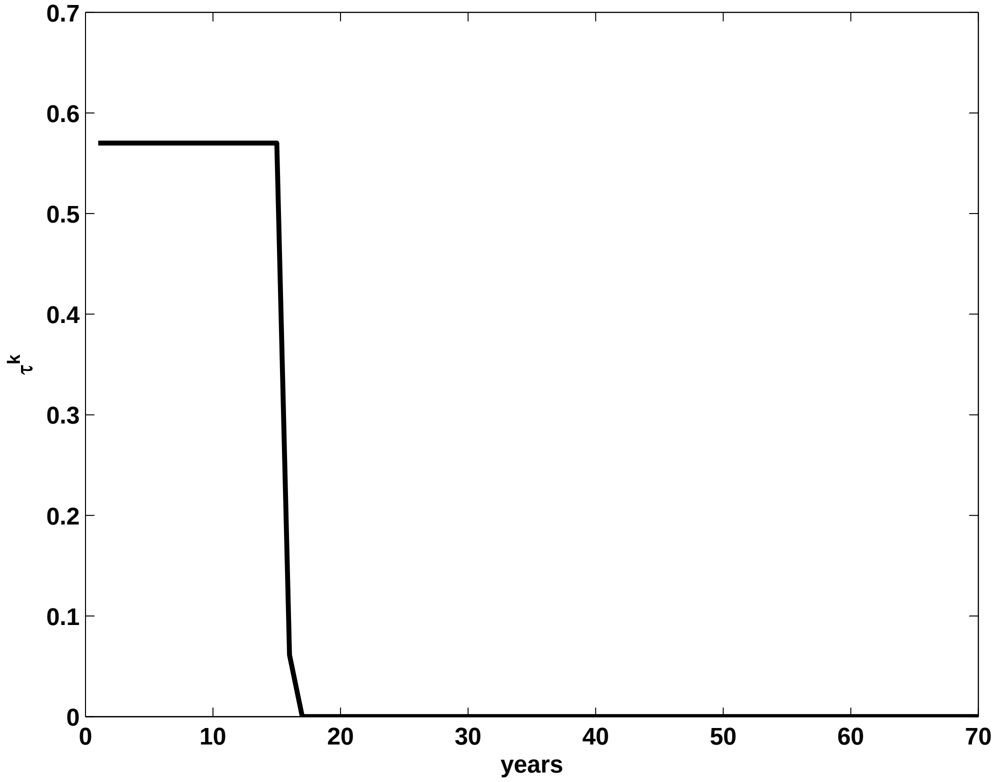


Figure 3: A typical time path for capital taxes

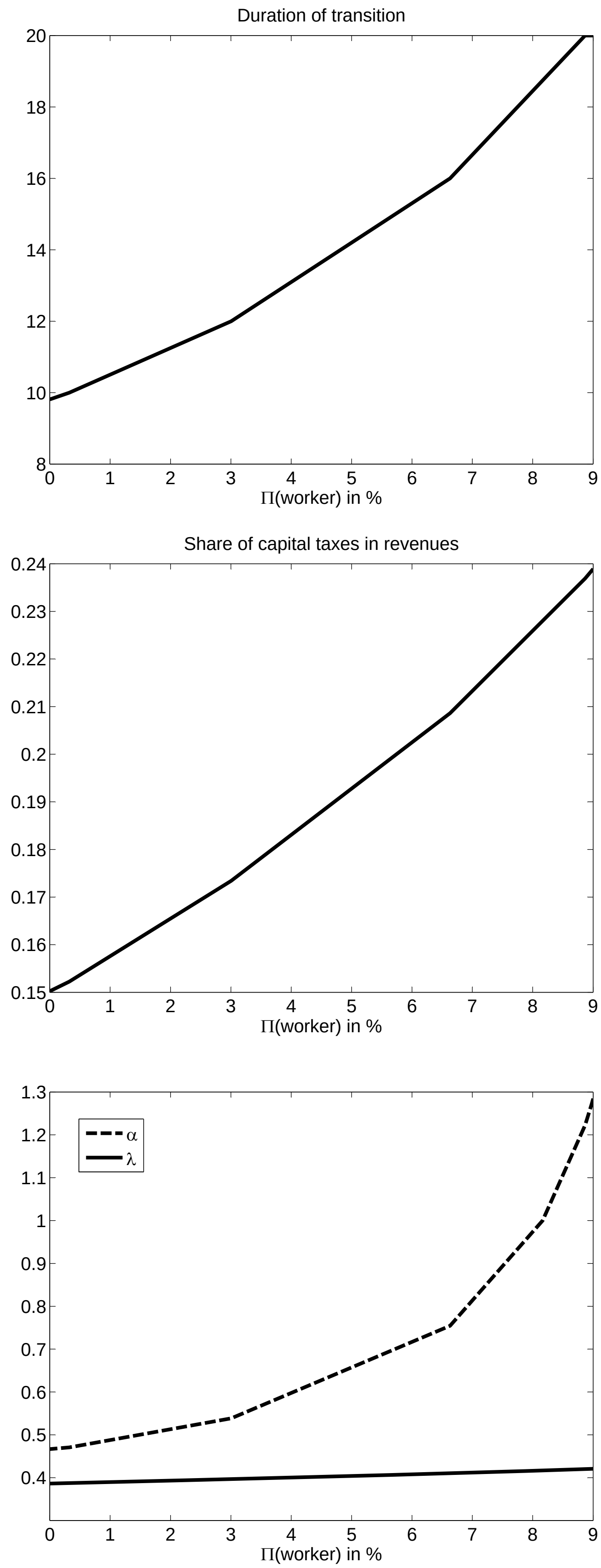


Figure 4: Properties of different POPI programs (baseline model)

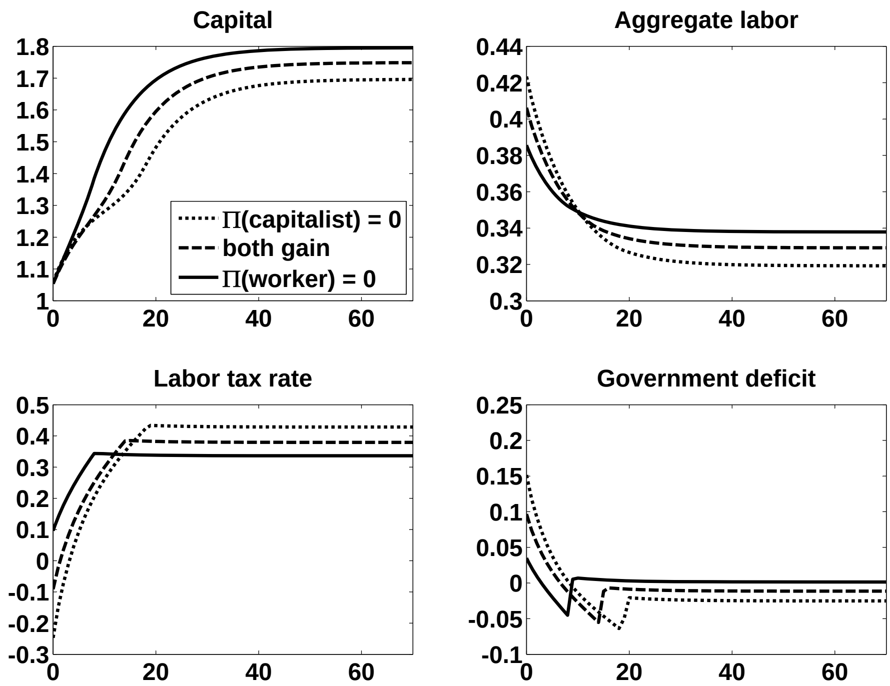
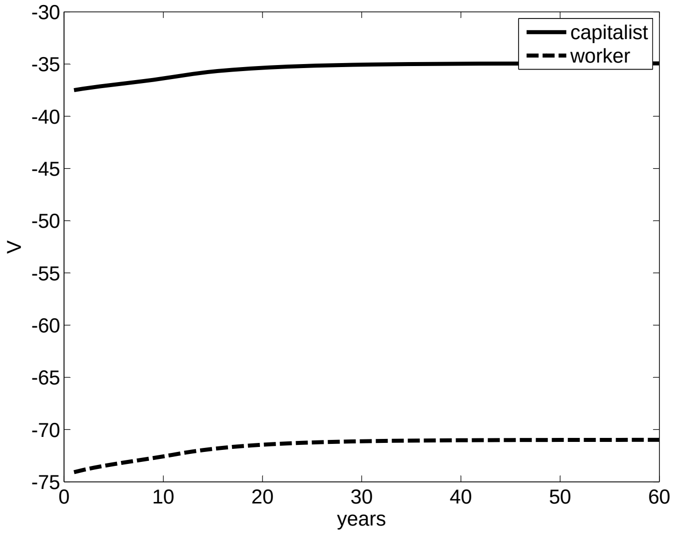


Figure 5: The time paths of selected variables for three POPI plans in the baseline model. (Time is in years.)



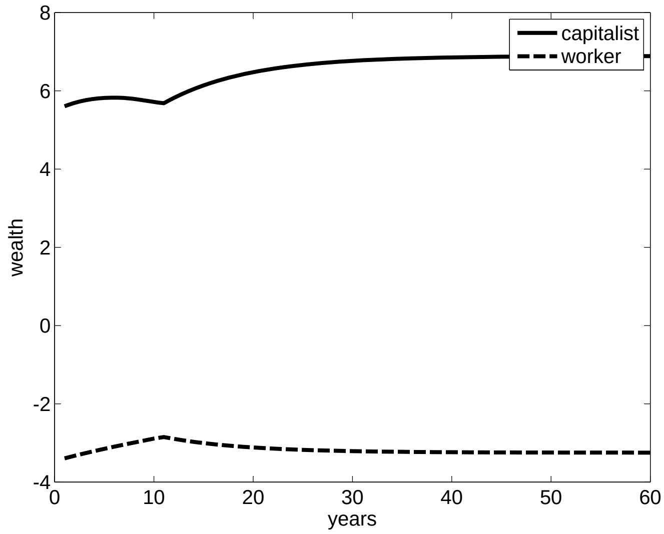


Figure 6: The evolution of agents' welfare and wealth over time (typical paths)