# Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks*

Sebastian Dyrda
University of Toronto

Marcelo Pedroni
University of Amsterdam

November 22, 2016

**Abstract**

In a standard incomplete markets model a Ramsey planner chooses time-varying paths of proportional capital and labor income taxes, lump-sum transfers (or taxes), and government debt. Distortive taxes reduce the variance cross-sectionally and over time of after-tax income, improving welfare for redistributive and insurance motives, which we quantify. Optimal capital income tax is higher than labor income tax in the long run; it provides insurance more efficiently. The government accumulates assets providing redistribution via general equilibrium price effects. The planners degree of inequality aversion only affects policy in the short run. Ignoring transition leads to significant welfare losses.

**Keywords:** Optimal Taxation; Heterogenous Agents; Incomplete markets
**JEL Codes:** E2; E6; H2; H3; D52

*Dyrda: University of Toronto, Max Gluskin House, 150 St. George Street, Toronto, ON M5S 3G7, Canada, sebastian.dyrda@utoronto.ca. Pedroni: University of Amsterdam, PO Box 15867, 1001 NJ Amsterdam, Netherlands, m.pedroni@uva.nl. Acknowledgements: We thank Anmol Bhandari, V.V. Chari, Carlos Eugênio da Costa, Jonathan Heathcote, Hans Holter, Larry Jones, Patrick Kehoe, Fabrizio Perri, Chris Phelan, José-Víctor Ríos-Rull, and Yikai Wang for many useful comments and discussions. We also thank the participants of the seminars at the University of Minnesota, Midwest Macro Meeting, NASM of the Econometric Society, NBER Summer Institute Macro and Public Finance, SED Tolouse, EEA-ESEM Geneva, Queen's University, University of Oslo, University of Amsterdam, Goethe University Frankfurt, Universidade Nova de Lisboa and University of Western Ontario. This work was carried out in part using computing resources at the University of Minnesota Supercomputing Institute.

How and to what extent should governments tax capital and labor income if they care about inequality and individual risk? This paper provides a quantitative answer to this question. We address it by solving a Ramsey problem in a general equilibrium model with heterogenous agents and uninsurable idiosyncratic labor income risk, originally developed and analyzed by Bewley (1986), Imrohoruglu (1989), Huggett (1993), and Aiyagari (1994), and from now on referred to as the standard incomplete markets (SIM) model.

The SIM model has been used extensively for positive analysis and been relatively successful at matching some basic facts about inequality and uncertainty[1]. In this environment agents face uncertainty with respect to their individual labor productivity which they cannot directly insure against (only a risk-free asset is available). Depending on their productivity realizations they make different savings choices which leads to endogenous wealth inequality. As a result, on top of the usual concern about not distorting agents decisions, a (utilitarian) Ramsey planner has two additional objectives: to redistribute resources across agents, and to provide insurance against their idiosyncratic productivity risk.

The study of optimal fiscal policy in the SIM model has largely focused, so far, on the maximization of steady state welfare[2]. In contrast, we allow policy to be *time varying* and the welfare function to depend on the associated *transition* path. We calibrate the initial steady state to replicate several aspects of the US economy; in particular the fiscal policy, the distribution of wealth, and statistical properties of the individual labor income process. The final steady state is, then, endogenously determined by the path of fiscal policy. As usual in the Ramsey literature, the planner finances an exogenous stream of government expenditures with the following instruments: proportional capital and labor income taxes and government debt. In contrast with the Ramsey literature, however, we allow for (possibly negative) lump-sum transfers. This would render the problem trivial in a representative-agent model, but that is not the case here.

We find that optimal capital income taxes are front-loaded hitting the imposed upper bound of $100$ percent for $33$ years then decreases to $45$ percent in the long run. Labor income taxes are reduced to less than half of their initial level, from $28$ percent to about $13$ percent in the long run. The ratio of lump-sum transfers to output is reduced to about a half of its initial level of $8$ percent and the government accumulates assets over time; the debt-to-output ratio decreases from $63$ percent to $-15$ percent in the long run (over the optimal transition asset level reaches $125$ percent of GDP though). Relative to keeping fiscal instruments at their initial levels, this leads to a welfare gain equivalent to a permanent $4.7$ percent increase in consumption.

[1]See examples of the calibration strategies in Domeij and Heathcote (2004) and Castañeda, Díaz-Giménez and Ríos-Rull (2003).

[2]See, for instance, Aiyagari and McGrattan (1998), Conesa, Kitao and Krueger (2009), and Nakajima (2010). Acikgoz (2015) and Hagedorn, Holter and Wang (2016) are notable exceptions.

Labor and capital income taxes are distortive, however, they are used to provide insurance and redistribution. The only uncertainty that agents face, in our environment, is with respect to their labor productivities[3]. Hence, labor income is the only risky part of the agents' income. By taxing it and rebating the extra revenue via lump-sum, the planner reduces the proportion of the agents' income that is uncertain and effectively provides insurance. On the other hand, capital income is particularly unequal and by taxing it the planner reduces the proportion of unequal income in total income and, this way, provides redistribution. To demonstrate exactly how the optimal policy reacts to changes in uncertainty and inequality we provide an analytic characterization of the solution to the Ramsey problem in a simple two-period version of the SIM model. In particular, we show that a higher intertemporal elasticity of substitution (Frisch elasticity) reduces the optimal capital (labor) income tax since it aggravates the distortions associated with it. The effect of government debt is more subtle. By decreasing debt the government crowds in capital which affects prices indirectly, in particular increasing wages and reducing interest rates which leads to a more uncertain but less unequal distribution of income. The optimal fiscal policy weighs all these effects against one another.

We decompose the average welfare gains of $4.7$ percent associated with implementing the optimal policy into three parts: (i) $3.1$ percent come from the more efficient allocation of aggregate resources due to the reduction of the distortions of agents' decisions; (ii) $6.4$ percent come from redistribution - the reduction in ex-ante inequality; and (iii) $-4.5$ percent come from the reduction in insurance - there is more uncertainty about individual consumption and labor streams under the optimal policy. The optimal policy implies an overall increase of capital taxes and a reduction of labor taxes. The net effect on the distortions of agents' savings and labor supply decisions is positive. The higher capital taxes decrease the proportion of the agents' income associated with the highly unequal asset income and lead to the redistributional gains. Finally, a lower labor income tax leads to a higher proportion of the agents' income being uncertain, thus the negative insurance effect that coincidentally almost exactly offsets the gains from the reduction in distortions.

We proceed to argue that disregarding transitory welfare effects in Ramsey problem can be severely misleading. To make this point we compute the stationary fiscal policy that maximizes welfare in the final steady state, which leads to a $9.8$ percent greater steady state welfare than the initial steady state. However, once transitory effects are considered, implementing this policy leads to a welfare *loss* of $6.4$ percent relative to keeping the initial fiscal policy. Relative to the fiscal policy that maximizes welfare over transition it leads to a welfare loss of $11.3$ percent.

In order to illustrate the role of market incompleteness and highlight why and how our results

[3] Panousi and Reis (2012) and Evans (2014) focus instead on investment risk. One justification for our focus on labor income risk is the fact that it is a bigger share of the total income for most agents in the economy. The bottom 80 percent in the distribution of net worth have a a share of labor income above 77 percent, in the 2007 SCF.

differ from the ones in the complete-markets Ramsey literature, we develop the following build-up. We start from the representative agent economy and sequentially introduce heterogeneity in initial assets; different (but constant and certain) individual productivity levels; and, finally, uninsurable idiosyncratic productivity risk which adds up to the SIM model. At each intermediate step, building on the work of Werning (2007), we analytically characterize and then numerically compute the optimal fiscal policy over transition identifying the effect of adding each feature. In particular, we show that the planner chooses to keep capital taxes at the upper bound in the initial periods if there is asset heterogeneity, before reducing it to zero. Productivity heterogeneity rationalizes positive (and virtually constant) labor taxes. The key qualitative difference of the solution once uninsurable idiosyncratic productivity risk is introduced is that long-run capital income taxes are set to a positive level, which therefore must have to do solely with the provision of insurance. One of the contributions of this paper is to quantify the optimal long-run capital taxes in the SIM model, which to our knowledge had not been done before.

Our benchmark results are for the utilitarian welfare function which implies a particular social choice with respect to the equality versus efficiency trade-off. We consider an alternative welfare function which allows us to control the degree of "inequality aversion" of the planner, i.e. the weight put on equality concerns. In particular we consider the case in which the planner completely disregards equality concerns. The optimal long-run levels of capital and labor taxes are surprisingly resistant to changes in the planner's inequality aversion. What does change significantly, however, is how long the capital tax is maintained at the upper bound; the more the planner "cares" about inequality the more years it keeps those taxes at the upper bound. When there are no equality concerns, taxes do not hit the upper bound for any periods making it clear that the only reason they do in the benchmark results is for redistributive purposes. Finally, we present robustness exercises with respect to the calibration of the labor income process and key elasticities, the magnitudes of the taxes are affected, but the qualitative features are maintained.

## Related Literature

This paper is related to several strands of literature. First, it is related to the literature on the steady state optimal fiscal policy in the SIM model. In an influential paper, Conesa, Kitao and Krueger (2009) solve for the tax system that maximizes steady state welfare in an overlapping generations SIM model. Their result includes an optimal long-run capital income tax of 36 percent. It is important to note that though this result is similar to ours the reasons behind it are different. They diagnose that their optimal capital tax level follows from the planner's inability to condition taxes on age, and the fact that a positive capital tax can mimic age-conditioned taxes in a welfare improving way (see Erosa and Gervais (2002)). This mechanism is not present in our analysis since we abstract from life-cycle issues.

Aiyagari (1995) and Chamley (2001) provide rationales for positive long-run capital taxes in environments similar to ours. Aiyagari (1995) shows that optimal taxes imply that the modified golden rule should hold in the long run which can only be achieved by taxing savings; the planner does not have precautionary motives while the agents do. We use this fact to corroborate our results and indeed the modified golden rule holds in the long run. Complementarily, Chamley (2001) shows, in a partial equilibrium version of the SIM model, that enough periods in the future every agent has the same probability of being in each of the possible individual (asset/productivity) states. It is, therefore, Pareto improving to transfer from the consumption-rich to the consumption-poor in the long run. If the correlation of asset holdings with consumption is positive, this transfer can be achieved by a positive capital tax rebated via lump-sum. In short, an agent's asset level in the long run is a good proxy for how lucky she has been; hence, taxing it is a good way to provide insurance in the long run. In recent work, Dávila, Hong, Krusell and Ríos-Rull (2012) solve the problem of a planner that is restricted to satisfy agents' budget constraints, but is allowed to choose the savings of each agent. If the consumption-poor's share of labor income is higher than the average, increasing the aggregate capital stock relative to the undistorted equilibrium can improve welfare through its indirect effect on wages and interest rates. In our setup, the Ramsey planner taxes capital to affect after tax interest rates directly and achieves the same goal.[4]

Using the SIM model, Aiyagari and McGrattan (1998) compute the level of debt-to-output that maximizes steady state welfare. Interestingly, they find that the optimal level is very close to the actual level in the data at that time, around $67$ percent. The fact that they abstract from the transitional dynamics makes the result even more remarkable: the government could chose its level of asset without having to finance it over time. It could, for instance, choose to have enough assets to finance all its expenditures and yet it chooses to remain in debt. By holding debt, the government crowds out capital increasing interest rates and decreasing wages. This effectively provides insurance since the proportion of uncertain labor income out of total income is reduced. This benefit is what drives the choice of the government to hold debt. However, there is another effect associated with such a policy; it increases inequality (the proportion of the unequal asset income out of total income increases). This negative effect is not particularly important in Aiyagari and McGrattan (1998) because their calibration focuses on matching labor income processes which leads to an underestimation of wealth inequality. Winter and Roehrs (2016) replicate their experiment with a calibration that targets wealth inequality statistics and find the opposite result, i.e. the government chooses to hold high levels of assets. Our calibration procedure is closer to that of Winter and Roehrs (2016), which elucidates our result that the Ramsey planner chooses to accumulate assets over time.

[4] The Online Appendix contains a more detailed discussion of the relationship between our results and theirs.

Heathcote, Storesletten and Violante (2014) and Gottardi, Kajii and Nakajima (2015) characterize the optimal fiscal policy in stylized versions of the SIM model. Their approaches lead to elegant and insightful closed-form solutions. The environment and Ramsey problem in Gottardi, Kajii and Nakajima (2015) is similar to ours except for the simplifications that yield tractability; i.e. exogenous labor supply, the absence of borrowing constraints, and i.i.d. shocks to human capital accumulation. Heathcote, Storesletten and Violante (2014), on the other hand, focus on different, though related, questions. By abstracting from capital accumulation, they are able to retain tractability in a model with progressive taxation, partial insurance, endogenous government expenditure and skill choices (with imperfect substitution between skill types). This leads to several interesting dimensions that, in our paper, we abstract from. However, the simplifications in these models do not allow them to match some aspects of the data, in particular the level of wealth inequality, which we find to be important for the determination of the optimal tax system.

We also contribute to the literature highlighting the importance of transition for policy prescriptions in incomplete markets models. Domeij and Heathcote (2004) use the SIM model to evaluate the implementation of a zero capital income tax policy taking into account the transitional welfare effects. They conclude that such a reform would be detrimental to welfare due to its transitory effect on inequality. Krueger and Ludwig (2013), Poschke, Kaymak and Bakis (2012), and Winter and Roehrs (2016) also conduct experiments in this spirit. Acikgoz (2015) and Hagedorn, Holter and Wang (2016) argue that the optimal long-run fiscal policy is independent of initial conditions and the transition towards it. They, then, proceed to study the properties of fiscal policy in the long run, but are silent about the optimal transition path which is the focus of this paper. For more on the relationship between their solution method and ours see Section 4.2.

There is an extensive literature that studies the Ramsey problem in complete-market economies with heterogeneity. The most well known result for the deterministic[5] subset of these economies is due to Judd (1985) and Chamley (1986): capital taxes should converge to zero in the long run. Among others, Jones, Manuelli and Rossi (1997) and Atkeson, Chari and Kehoe (1999) show this result is robust to a relaxation of a number of assumptions. Werning (2007) characterizes optimal policy for this class of economies using the same set of fiscal instruments that we use, in particular, allowing for lump-sum transfers or taxes. Laczo, Marcet and Greulich (2015) characterize and compute the Pareto improving capital and labor income taxes while addressing the criticism by Straub and Werning (2014) imposing a bound on agents consumption levels. In Section 5 we characterize analytically fiscal policy over the optimal transitions and link the results to these

[5] Aiyagari, Marcet, Sargent and Seppala (2002) consider market incompleteness with respect to aggregate risk without heterogeneity. Bassetto (2014) studies how fiscal policy should respond to aggregate fluctuations in a complete-market model with heterogeneous agents. Bhandari, Evans, Golosov and Sargent (2016) consider both market incompleteness and heterogeneity.

studies.

The New Dynamic Public Finance literature takes an alternative approach to answer our initial question. It focuses on the design of a mechanism that would allow the planner to extract information about the agents' unobservable productivities efficiently. It assumes tax instruments are unrestricted and in this sense it dominates the Ramsey approach in terms of generality, since the latter ignores the information extraction problem[6] and imposes ad-hoc linearity restrictions on the tax system. One of the main results steaming from this literature is the inverse Euler equation; see Golosov, Kocherlakota and Tsyvinski (2003). Farhi and Werning (2012) show that starting from the allocations from the steady state of an undistorted SIM model and applying perturbations to implement the inverse Euler equation leads to small welfare gains, of the order of $0.2$ percent. Moreover, it is difficult to solve the private information problem in dynamic economies with persistent shocks. Farhi and Werning (2013) and Troshkin, Tsyvinski and Golosov (2010) have made advancements in this direction in partial equilibrium settings without capital and find that restrictions to linear taxes lead to small welfare losses. Our view is that, even if only as a benchmark to more elaborate tax systems, it is useful to understand the properties of a simpler optimal linear tax system in a quantitative general equilibrium environment.

The rest of the paper is organized as follows. Section 1 illustrates the main mechanism behind our results in a two-period economy. Section 2 describes the infinite horizon model, sets up the Ramsey problem and discusses our solution technique. Section 3 describes the calibration. Section 4 presents the main results of the paper. Section 5 presents the build-up from the complete market economy results to our main results. Sections 6 and 7 provide results for alternative welfare functions and calibrations and Section 8 concludes.

# 1 Mechanism: Two-Period Economy

In the SIM model, there are two dimensions of heterogeneity: productivity and wealth. Agents have different levels of productivity which follow an exogenous stochastic process. In addition, markets are incomplete and only a risk-free asset exists. Therefore, the idiosyncratic productivity risk cannot be diversified away. It follows that the history of shocks, affects the amount of wealth accumulated by each agent and there is an endogenously determined distribution of wealth.

In a two-period economy, it is possible to evaluate how each dimension of heterogeneity affects the optimal tax system. Since there is no previous history of shocks, the initial wealth inequality can be set exogenously. In this section, we characterize, under some assumptions about preferences, the optimal tax system when the government has access to linear labor and capital income taxes, and lump-sum transfers. The lump-sum transfers are allowed to be negative, and the govern-

[6]The Ramsey planner is also unable to observe productivity levels, it is not allowed to condition taxes on them.

ment could finance all necessary revenue with this non-distorvive instrument. In this section we explain why it chooses to do otherwise. First, we assume agents have the same level of wealth but face an idiosyncratic productivity shock - we call this the *uncertainty economy* . Then, we shut down uncertainty and introduce ex-ante wealth inequality - this is referred to as the *inequality economy* . Next we discuss the relationship with the infinite horizon problem.[7]

## 1.1 Uncertainty economy

Consider an economy with a measure one of ex-ante identical agents who live for two periods. Suppose they have time-additive, von Neumann-Morgenstern utility functions. Denote the period utility function by $u(c,n)$ where $c$ and $n$ are the levels of consumption and labor supplied. Assume $u$ satisfies the usual conditions and denote the discount factor by $\beta$. In the first period each agent is endowed with $\omega$ units of the consumption good which can be either consumed or invested into a risk-free asset, $a$, and supplies $\bar{n}$ units of labor inelastically.

In period 2, consumers receive income from the asset they saved in period 1 and from labor. Labor is supplied endogenously by each agent in period 2 and the individual labor productivity, $e$, is random and can take two values: $e_L$ with probability $\pi$ and $e_H > e_L$ with probability $1-\pi$, with the normalization $\pi e_L + (1-\pi) e_H = 1$. Due to the independence of shocks across consumers a law of large numbers operates so that in period 2 the fraction of agents with $e_L$ is $\pi$ and with $e_H$ is $(1-\pi)$. Letting $n_i$ be the labor supply of an agent with productivity $e_i$, it follows that the aggregate labor supply is $N = \pi e_L n_L + (1-\pi) e_H n_H$.

The planner needs to finance an expenditure of $G$ in period 2. It has three instruments available: labor and capital income taxes, $\tau^n$ and $\tau^k$, and lump-sum transfers $T$ which can be positive or negative. Let $w$ be the wage rate and $r$ the interest rate. The total period 2 income of an agent with productivity $e_i$ is, therefore, $(1-\tau^n) w e_i n_i + \left(1 + \left(1-\tau^k\right) r\right) a + T$. In period 2, output is produced using capital, $K$, and labor and a constant-returns-to-scale neoclassical production function $f(K,N)$. We assume that $f(\cdot)$ is net of depreciation.

**Definition 1** *A* ***tax distorted competitive equilibrium*** *is a vector* $(K, n_L, n_H, r, w; \tau^n, \tau^k, T)$ *such that*

1. $(K, n_L, n_H)$ *solves*

$$\max_{a, n_L, n_H} u(\omega - a, \bar{n}) + \beta E\left[u(c_i, n_i)\right] \quad \text{s.t. } c_i = (1-\tau^n) w e_i n_i + \left(1 + \left(1-\tau^k\right) r\right) a + T;$$

2. $r = f_K(K,N)$, $w = f_N(K,N)$, *where* $N = \pi e_L n_L + (1-\pi) e_H n_H$;

[7] The Online Appendix discusses the case which there is uncertainty *and* inequality and the connection between the results obtained here and the ones in Dávila et al. (2012).

*3. and,* $\tau^n wN + \tau^k rK = G + T$.

The Ramsey problem is to choose $\tau^n$, $\tau^k$, and $T$ to maximize welfare. Since agents are ex-ante identical there is no ambiguity about which welfare function to use, it is the expected utility of the agents. If there is no risk, i.e. $e_L = e_H$, the agents are also ex-post identical and the usual representative agent result applies: since negative lump-sum transfers are available, it is optimal to obtain all revenue via this undistortive instrument and set $\tau^n = \tau^k = 0$.

In order to provide a sharp characterization of the optimal tax system we make the following assumption discussed below[8].

**Assumption 1** *No income effects on labor supply and constant Frisch elasticity, $\kappa$, i.e.*

$$u_{cn} - u_{cc}\frac{u_n}{u_c} = 0, \quad \textit{and} \quad \frac{u_{cc}u_n}{n\left(u_{cc}u_{nn} - u_{cn}^2\right)} = \kappa.$$

We pursue a variational approach. Suppose $\left(K, n_L, n_H, r, w; \tau^n, \tau^k, T\right)$ is a tax distorted equilibrium[9]. Consider a small variation on the tax system $\left(d\tau^n, d\tau^k, dT\right)$, such that all the equilibrium conditions are satisfied. Then, evaluate the effect of such a variation on welfare, taking as given the optimal decision rules of the agents. Using this method we establish the following proposition.

**Proposition 1** *In the uncertainty economy, if $u$ satisfies Assumption $A$, then, the optimal tax system is such that $\tau^k = 0$,*

$$\tau^n = \frac{(\nu - 1)\,\pi(1 - \pi)\,(e_H n_H - e_L n_L)}{(\nu - 1)\,\pi(1 - \pi)\,(e_H n_H - e_L n_L) + \kappa N\,(\pi\nu + (1 - \pi))} > 0, \tag{1.1}$$

*where $\nu \equiv \frac{u_c(c_L, n_L)}{u_c(c_H, n_H)}$, and $T < 0$ balances the budget.*

**Proof.** See Appendix A.1. ■

Notice that the planner could choose to finance $G$ with $T$ but chooses a positive distortive labor income tax instead. The revenue from labor taxation is rebated via lump-sum transfers and the proportion of the agents' income that comes from the uncertain labor income is reduced. Hence, this tax system effectively provides insurance to the agents. Why not provide full insurance by taxing away all the labor income? This is exactly what would happen if labor were supplied

[8]In a similar two-period environment, Gottardi et al. (2014) characterize the solution to Ramsey problem without Assumption A. However, they impose an alternative assumption about the sign of general equilibrium effects, which are satisfied under Assumption A. Further, this assumption allows us to provide a sharper characterization of the optimal tax system (besides the signs of taxes we also characterize the levels).

[9]Since the equilibrium does not exist for $\tau^n \geq 1$ or $\tau^k \geq (1 + r)/r$, we impose the restrictions that $\tau^n < 1$ and $\tau^k < (1 + r)/r$.

inelastically. In fact, notice that in this case $\kappa = 0$ and equation (1.1) implies $\tau^n = 1$. However, with an endogenous labor supply the planner has to balance two objectives: minimize distortions to agents' decisions and provide insurance. This balance is explicit in equation (1.1) seeing as a higher $\kappa$ implies a lower $\tau^n$. That is, the more responsive labor supply is to changes in labor taxes the more distortive these taxes are and the planner chooses a lower labor tax. In the limit, if $\kappa \to \infty$ it will be optimal to set $\tau^n = 0$.

With income effects on labor supply, distortions of the savings decision would spill over to the labor supply decision and vice-versa. Thus, it could be optimal, for instance, to choose $\tau^k$ so as to mitigate the distortion imposed by a positive $\tau^n$. This complex relationship complicates the analysis considerably. Assumption 1 unties this relationship and as a result it is optimal to set $\tau^k = 0$.

Next, suppose that $e_L = 1 - \epsilon^{unc}/\pi$ and $e_H = 1 + \epsilon^{unc}/(1-\pi)$, so that $\epsilon^{unc}$ is a mean preserving spread on the productivity levels. It is easy to see that if $\epsilon^{unc} = 0$ equation (1.1) implies that $\tau^n = 0$. The effect of an increase in $\epsilon^{unc}$ on the optimal $\tau^n$ is not as obvious since the right hand side of equation (1.1) contains endogenous variables. An application of the implicit function theorem, however, clarifies that as long as $\partial\nu/\partial\epsilon^{unc} > 0$ and $\partial\nu/\partial\tau^n < 0$, it follows that $\partial\tau^n/\partial\epsilon^{unc} > 0$, i.e. the optimal labor income tax is increasing in the level of risk in the economy. Under standard calibrations, the equilibrium ratio of marginal utilities, $\nu$, is in fact increasing in the level of risk ($\partial\nu/\partial\epsilon^{unc} > 0$) and decreasing in the labor income tax ($\partial\nu/\partial\tau^n < 0$).

## 1.2 Inequality economy

Consider the environment described above only without uncertainty and with initial wealth inequality. That is, suppose the productivity levels do not vary between agents, i.e. $e_L = e_H = 1$, and that $\omega$ can take two values: $\omega_L$ for a proportion $p$ of the agents and $\omega_H > \omega_L$ for the rest, with $\bar{\omega} \equiv p\omega_L + (1-p)\omega_H$.

**Definition 2** *A* ***tax distorted competitive equilibrium*** *is* $\left(a_L, a_H, n_L, n_H, r, w; \tau^n, \tau^k, T\right)$ *such that*

1. *For* $i \in \{L, H\}$, $(a_i, n_i)$ *solves*

$$\max_{a_i, n_i} u\left(\omega_i - a_i, \bar{n}\right) + \beta u\left(c_i, n_i\right), \quad \textit{s.t. } c_i = \left(1-\tau^n\right) w n_i + \left(1 + \left(1-\tau^k\right) r\right) a_i + T;$$

2. $r = f_K(K, N)$, $w = f_N(K, N)$, *where* $K = pa_L + (1-p)a_H$ *and* $N = pn_L + (1-p)n_H$;

3. *and,* $\tau^n w N + \tau^k r K = G + T$.

In this economy the concept of optimality is no longer unambiguous. Since agents are different ex-ante, a decision must be made with respect to the social welfare function. In what follows, by optimal we mean the one that maximizes $W \equiv pU_L + (1-p)\, U_H$; the utilitarian welfare function. The following proposition follows.

**Proposition 2** *In the inequality economy, if $u$ satisfies Assumption $A$ and has CARA or is GHH, as in equation (3.1), then the optimal tax system is such that $\tau^n = 0$,*

$$\tau^k = \frac{\left(\frac{1+r}{r}\right)(\nu-1)\, p(1-p)\,(\omega_H-\omega_L)}{(\nu-1)\, p(1-p)\,(\omega_H-\omega_L) + \frac{\rho}{\psi}\,(p\nu+(1-p))} > 0, \tag{1.2}$$

*where $\rho \equiv \frac{2+(1-\tau^k)r}{2+r}$ for CARA, $\rho \equiv \frac{1+\beta^{-\frac{1}{\sigma}}(1+(1-\tau^k)r)^{\frac{\sigma-1}{\sigma}}}{1+r+\beta^{\frac{1}{\sigma}}(1+(1-\tau^k)r)^{\frac{1}{\sigma}}}$ for GHH, and $\psi$ is the level of absolute risk aversion*[10]. *$T<0$ balances the budget.*

**Proof.** See Appendix A.2. ■

The planner chooses a positive capital income tax which distorts savings decisions but allows for redistribution between agents. The ex-ante wealth inequality is exogenously given. However, agents with different wealth levels in the first period will save different amounts and have different asset levels in the second period. This endogenously generated asset inequality is the one the tax system is able to affect. A positive capital tax rebated via lump-sum transfers directly reduces the proportion of the agents' income that will be dependent on unequal asset income achieving the desired redistribution which implies a reduction of consumption inequality (by assumption, there is no labor supply inequality).

One of the key elements of equation (1.2) is the inverse of the coefficient of absolute risk aversion, $1/\psi$, which is proportional to the agents' intertemporal elasticity of substitution. This elasticity indicates the responsiveness of savings to changes in $\tau^k$. Hence, the higher this elasticity is the lower is the optimal level of $\tau^k$, since providing redistribution becomes more costly. The $\tau^n = 0$ result is again associated with Assumption 1.

Assuming that $\omega_L = 1-\epsilon^{ine}/p$ and $\omega_H = 1-\epsilon^{ine}/\,(1-p)$, the effect of an increase in the mean preserving spread, $\epsilon^{ine}$, on the optimal $\tau^k$ can again be found by applying the implicit function theorem on equation (1.2). It follows that, if $\partial\nu/\partial\epsilon^{ine} > 0$ and $\partial\nu/\partial\tau^k < 0$, then $\partial\tau^k/\partial\epsilon^{ine} > 0$; the optimal capital income tax is increasing in the level of inequality in the economy. If $u$ satisfies Assumption $A$ and has CARA one can show that this is always the case.

## 1.3 Relationship with infinite horizon problem

The two-period examples are useful to understand some of the key trade-offs faced by the Ramsey planner, since they allow for the exogenous setting of the levels of uncertainty (ex-post risk) and inequality (ex-ante risk). In the infinite horizon version of the SIM model, however, these concepts

[10]The level of absolute risk aversion is endogenous is the GHH case.

are inevitably intertwined. The characterization of the optimal tax system, therefore, becomes considerably more complex. Labor income taxes affect not only the level of uncertainty through the mechanism described above, but also the labor income inequality and the distribution of assets over time. An agent's asset level at a particular period depends not only on its initial value, but on the history of shocks this agent has experienced. Therefore, capital income taxation affects not only the ex-ante risk faced by the agents but also the ex-post. Nevertheless, these results are useful to understand some features of the optimal fiscal policy in the infinite horizon model as will become clear in what follows. In particular, Section 7.2 shows that the comparative statics with respect to agents' intertemporal elasticity of substitution and Frisch elasticity described in this section are also pertinent for the infinite horizon problem.

# 2 The Infinite-Horizon Model

Time is discrete and infinite, indexed by $t$. There is a continuum of agents with standard preferences $E_0\left[\sum_t \beta^t u\left(c_t, n_t\right)\right]$ where $c_t$ and $n_t$ denote consumption and labor supplied in period $t$ and $u$ satisfies the usual conditions. Individual labor productivity, $e \in E$ where $E \equiv \{e_1, ..., e_L\}$, are i.i.d. across agents and follow a Markov process governed by $\Gamma$, a transition matrix[11]. Agents can only accumulate a risk-free asset, $a$. Let $A \equiv [\underline{a}, \infty)$ be the set of possible values for $a$ and $S \equiv E \times A$. Individual agents are indexed by the a pair $(e, a) \in S$. Given a sequence of prices $\{r_t, w_t\}_{t=0}^{\infty}$, labor income $\{\tau_t^n\}_{t=0}^{\infty}$, (positive) capital income $\{\tau_t^k\}_{t=0}^{\infty}$, and lump-sum transfers $\{T_t\}_{t=0}^{\infty}$, each household, at time $t$, chooses $c_t\left(a, e\right)$, $n_t\left(a, e\right)$, and $a_{t+1}\left(a, e\right)$ to solve

$$v_t(a, e) = \max u(c_t(a, e), n_t(a, e)) + \beta \sum_{e_{t+1} \in E} v_{t+1}(a_{t+1}(a, e), e_{t+1})\Gamma_{e, e_{t+1}}$$

subject to

$$\begin{aligned}(1 + \tau^c)c_t(a, e) + a_{t+1}(a, e) = \left(1 - \tau_t^n\right) w_t e n_t(a, e) + (1 + (1 - I_{\{a \geq 0\}}\tau_t^k)r_t)a + T_t \\ a_{t+1}(a, e) \geq \underline{a}.\end{aligned}$$

Note that value and policy functions are indexed by time, because policies $\{\tau_t^k, \tau_t^n, T_t\}_{t=0}^{\infty}$ and aggregate prices $\{r_t, w_t\}_{t=0}^{\infty}$ are time-varying. The consumption tax, $\tau^c$, is a parameter[12]. Let

[11] A law of large numbers operates so that the probability distribution over $E$ at any date $t$ is represented by a vector $p_t \in \mathbb{R}^L$ such that given an initial distribution $p_0$, $p_t = p_0\Gamma^t$. In our exercise we make sure that $\Gamma$ is such that there exists a unique $p^* = \lim_{t\to\infty} p_t$. We normalize $\sum_i p_i^* e_i = 1$.

[12] It is not without loss of generality that we do not allow the planner to choose $\tau_c$. There are two reasons for this choice. The first is practical, we are already using the limit of the computational power available to us, and allowing for one more choice variable would increase it substantially. Second, in the US capital and labor income taxes are chosen by the Federal Government while consumption taxes are chosen by the states, so this Ramsey problem can be understood as the one relevant for a Federal Government. We add $\tau^c$ as a parameter for calibration purposes.

$\{\lambda_t\}$ be a sequence of probability measures over the Borel sets $\mathcal{S}$ of $S$ with $\lambda_0$ given. Since the path for taxes is known, there will be a deterministic path for prices and for $\{\lambda_t\}_{t=0}^{\infty}$. Hence, we do not need to keep track of the distribution as an additional state; time is a sufficient statistic.

Competitive firms own a constant-returns-to-scale technology $f(\cdot)$ that uses capital, $K_t$, and efficient units of labor, $N_t$, to produce output each period ($f(\cdot)$ denotes output net of depreciation, $\delta$ denotes the depreciation rate). A representative firm exists that solves the usual static problem. The government needs to finance an exogenous constant stream of expenditure, $G$, and lump-sum transfers with taxes on consumption, labor income, and (positive) capital income. It can also issue debt $\{B_{t+1}\}$ and, thus, has the following intertemporal budget constraint

$$G + r_t B_t = B_{t+1} - B_t + \tau^c C_t + \tau_t^n w_t N_t + \tau_t^k r_t \hat{A}_t - T_t, \tag{2.1}$$

where $C_t$ is aggregate consumption and $\hat{A}_t$ is the tax base for the capital income tax.

**Definition 3** *Given $K_0$, $B_0$, an initial distribution $\lambda_0$ and a policy $\pi \equiv \{\tau_t^k, \tau_t^n, T_t\}_{t=0}^{\infty}$, a **competitive equilibrium** is a sequence of value functions $\{v_t\}_{t=0}^{\infty}$, an allocation $X \equiv \{c_t, n_t, a_{t+1}, K_{t+1}, N_t, B_{t+1}\}_{t=0}^{\infty}$, a price system $P \equiv \{r_t, w_t\}_{t=0}^{\infty}$, and a sequence of distributions $\{\lambda_t\}_{t=1}^{\infty}$, such that for all $t$:*

1. *Given $P$ and $\pi$, $c_t(a,e)$, $n_t(a,e)$, and $a_{t+1}(a,e)$ solve the household's problem and $v_t(a,e)$ is the respective value function;*

2. *Factor prices are set competitively,*

$$r_t = f_K(K_t, N_t), \quad w_t = f_N(K_t, N_t);$$

3. *The probability measure $\lambda_t$ satisfies*

$$\lambda_{t+1} = \int_S Q_t\left((a,e), \mathcal{S}\right) d\lambda_t, \quad \forall \mathcal{S} \in S$$

   *where $Q_t$ is the transition probability measure;*

4. *The government budget constraint, (2.1), holds and debt is bounded;*

5. *Markets clear,*

$$C_t + G_t + K_{t+1} - K_t = f\left(K_t, N_t\right), \quad K_t + B_t = \int_{A\times E} a_t(a,e) d\lambda_t.$$

## 2.1 The Ramsey Problem

We now turn to the problem of choosing the optimal tax policy in the economy described above. We assume that, in period $0$, the government announces and commits to a sequence of future taxes $\{\tau_t^k, \tau_t^n, T_t\}_{t=1}^{\infty}$, taking period $0$ taxes as given. We need the following definitions:

**Definition 4** *Given $K_0$, $B_0$, $\lambda_0$, and $\{\tau_0^k, \tau_0^n, T_0\}$, for every policy $\pi$, **equilibrium allocation rules** $X(\pi)$ and **equilibrium price rules** $P(\pi)$ are such that $\pi$, $X(\pi)$, $P(\pi)$ and corresponding $\{v_t\}_{t=0}^{\infty}$ and $\{\lambda_t\}_{t=1}^{\infty}$ constitute a competitive equilibrium.*

**Definition 5** *Given $K_0$, $B_0$, $\lambda_0$, and $\{\tau_0^k, \tau_0^n, T_0\}$, and a welfare function $W(\pi)$, the **Ramsey problem** is to $\max_\pi W(\pi)$ such that $X(\pi)$ and $P(\pi)$ are equilibrium allocation and price rules.*

In our benchmark experiments we assume that the Ramsey planner maximizes the utilitarian welfare function: the ex-ante expected lifetime utility of a newborn agent who has its initial state, $(a, e)$, chosen at random from the initial stationary distribution $\lambda_0$. The planner's objective is, thus, given by

$$W(\pi) = \int_S E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(a, e|\pi\right), n_t\left(a, e|\pi\right)\right) d\lambda_0.$$

In Section 7 we provide results for alternative welfare functions.

## 2.2 Solution method

We solve this problem numerically. Given an initial stationary equilibrium, for any policy $\pi$ we can compute the transition to a new stationary equilibrium consistent with the policy, as long as the taxes become constant at some point. and calculate welfare $W(\pi)$. We then search for the policy $\pi$ that maximizes $W(\pi)$. This is, however, a daunting task since it involves searching in the space of infinite sequences. In order to make it computationally feasible we impose the following ad-hoc constraints: that each path $\{\tau_t^k, \tau_t^n, T_t\}_{t=1}^{\infty}$ is smooth over time and become constant after a finite number of periods. We denote the set of policies that satisfy these properties by $\Pi_R$. These conditions are restrictive, but they allow the problem to be solved and are flexible enough to characterize some of the key features of the optimal paths of taxes.

The statement about the ad-hoc constraints must be qualified. In Section 5 we show that in complete markets economies optimal capital taxes should be front-loaded. Hence, in defining the set $\Pi_R$ we take this under consideration. That is, we allow capital taxes to hit the imposed upper bound of $100$ percent for the first $t^*$ periods, where a model period is equivalent to one calendar year. Importantly, $t^*$ is endogenously chosen and is allowed to be zero, so the fact that the solution displays a capital tax at the upper bound for a positive amount of periods is not an assumption but a result. Other than this, we assume that the paths for $\left\{\tau_t^k\right\}_{t=t^*+1}^{\infty}$ and $\left\{\tau_t^n, T_t\right\}_{t=1}^{\infty}$ follow splines with nodes set at exogenously selected periods. The placement of the nodes is arbitrary, we started with a small number of them and sequentially added more until the solution converged. In the main experiment the planner was allowed to choose $15$ variables[13] in total: $t^*$, $\tau_{t^*+1}^k$, $\tau_{45}^k$, $\tau_{60}^k$, $\tau_{100}^k$, $\tau_1^n$, $\tau_{15}^n$, $\tau_{t^*+1}^n$, $\tau_{45}^k$, $\tau_{60}^k$, $\tau_{100}^k$, $T_1$, $T_{15}$, $T_{t^*+1}$, $T_{45}$, $T_{60}$, and $T_{100}$. In the Online Appendix we include figures that compare the optimal fiscal policy computed with $4$, $6$, $9$, $12$, and $15$ variables; the welfare gains change from $4.73$ to $4.74$ in the last step. In the intermediate periods the paths follow a spline function and after the final period they become constant at the last level. The choice of the periods $1$, $15$, $45$, $60$, and $100$, were a result of the fact that for experiments with less nodes, the optimal $t^*$ was always close to $30$, hence we placed the nodes at the same distance from each other except for the last ones which are supposed to capture the long run levels[14].

Solving the problem described above is a particularly hard computational task. Effectively we are maximizing $W(\pi)$ on the domain $\pi \in \Pi_R$, where each element of $\Pi_R$ can be defined by a vector with a finite number of elements (the nodes described above). We know very little about its properties; it is a multivariate function with potentially many kinks, irregularities and multiple local optima. Thus, we need a powerful and thorough procedure to make sure we find the global optimum. We use a global optimization algorithm that randomly draws a very large number of policies in $\Pi_R$ and computes the transition between the exogenously given initial stationary equilibrium and a final stationary equilibrium that depends on the policy. Then, we compute welfare $W(\pi)$ for each of those policies and select those that yield the highest levels of welfare. These selected policies are then clustered, similar policies placed in the same cluster. For each cluster we run an efficient derivative free local optimizer. The whole procedure is repeated depending on how many local optima have been found and a Bayesian stopping rule is used to figure out if enough global procedures have been run. A more detailed description of the algorithm can be found in the Online Appendix[15].

# 3 Calibration

We calibrate the initial stationary equilibrium of the model economy to replicate key properties of the US economy relevant for the shape of the optimal fiscal policy. Table 1 summarizes our

[13] We combine $t^*$ and $\tau_{t^*+1}^k$ into one variable and one of the final taxes must be chosen so that government debt is bounded.

[14] If the solver chooses $t^*$ close to one of these predetermined nodes the algorithm replaces that node for $t = 30$. For instance, if $t^* = 43$ the periods become $1$, $15$, $30$, $t^* + 1$, $60$, and $100$.

[15] The algorithm was parallelized for multiple cores. For each global iteration, we drew $147,456$ policies and computed the transition and welfare for each of them. The number of transitions run for each cluster is endogenously determined by the local solver, on average it amounted to around $150$ transitions to find each local maximum. A total of $8$ global iterations were needed. We performed our analysis on the Mesabi cluster at the Minnesota Supercomputing Institute using $1152$ cores.

parameter choices together with the targets we use to discipline their values and their model counterparts. We use data from the NIPA tables for the period between 1995 and 2007[16] and from the 2007 Survey of Consumer Finances (SCF).

Table 1: Benchmark Model Economy: Target Statistics and Parameters

| **Statistic** | **Target** | **Model** | **Parameter** | **Value** |
|---|---|---|---|---|
| **Preferences and Technology** | | | | |
| Intertemporal elast. of substitution | 0.50 | 0.50 | $\sigma$ | 2.000* |
| Frisch elasticity | 0.72 | 0.72 | $\nu$ | 0.720* |
| Average hours worked | 0.30 | 0.30 | $\chi$ | 4.120 |
| Capital to output | 2.72 | 2.71 | $\beta$ | 0.965 |
| Capital income share | 0.38 | 0.38 | $\alpha$ | 0.380* |
| Investment to output | 0.27 | 0.27 | $\delta$ | 0.100 |
| **Borrowing Constraint** | | | | |
| % of hhs with wealth $< 0$ | 18.6 | 19.1 | $\underline{a}/Y$ | $-0.034$ |
| **Fiscal Policy** | | | | |
| Capital income tax (%) | 36.0 | 36.0 | $\tau_k$ | 0.360* |
| Labor income tax (%) | 28.0 | 28.0 | $\tau_n$ | 0.280* |
| Consumption tax (%) | 5.0 | 5.0 | $\tau_c$ | 0.050* |
| Transfer to output (%) | 8.0 | 8.0 | $T/Y$ | 0.080 |
| Debt-to-output (%) | 63.0 | 63.0 | $G/Y$ | 0.146 |
| **Labor Productivity Process** | | | | |
| Wealth Gini index | 0.82 | 0.81 | $e_1/e_2$ | 0.625 |
| % of wealth in 1st quintile | $-0.2$ | $-0.2$ | $e_3/e_2$ | 3.900 |
| % of wealth in 4th quintile | 11.2 | 10.2 | $\Gamma_{11}$ | 0.956 |
| % of wealth in 5th quintile | 83.4 | 83.4 | $\Gamma_{12}$ | 0.043 |
| % of wealth in top 5% | 60.3 | 60.8 | $\Gamma_{21}$ | 0.071 |
| Corr. btw wealth and labor income | 0.29 | 0.29 | $\Gamma_{22}$ | 0.929 |
| Autocorr. of labor income | 0.90 | 0.90 | $\Gamma_{31}$ | 0.012 |
| Std of labor income | 0.20 | 0.20 | $\Gamma_{32}$ | 0.051 |

Notes: Parameter values marked with (*) were set exogenously, all the others were endogenously and jointly determined.

[16]We choose this time period to be consistent with the one used to pin down fiscal policy parameters which we take from Trabandt and Uhlig (2011) and also to prevent the Great Recession to affect our results.

## 3.1 Preferences and technology

We assume GHH preferences (see Greenwood et al. (1988)) with period utility given by

$$u(c, n) = \frac{1}{1-\sigma}\left(c - \chi\frac{n^{1+\frac{1}{\kappa}}}{1+\frac{1}{\kappa}}\right)^{1-\sigma}, \tag{3.1}$$

where $\sigma$ is the coefficient of relative risk aversion, $\kappa$ is the Frisch elasticity of labor supply and $\chi$ is the weight on the disutility of labor. These preferences exhibit no wealth effects on labor supply, which is consistent with some microeconometric evidence showing these effects are in fact small. See Holtz-Eakin et al. (1993), Imbens et al. (2001) and Chetty et al. (2012) for details.[17]

Further, they imply that aggregate labor supply is independent of the distribution of wealth which is convenient for computing out of steady state allocations in our main experiment. We set the intertemporal elasticity of substitution to $0.5$; the number frequently used in the literature (e.g. Dávila et al. (2012) and Conesa et al. (2009)). For the Frisch elasticity, $\kappa$, we rely on estimates from Heathcote et al. (2010) and use $0.72$. This value is intended to capture both the intensive and the extensive margins of labor supply adjustment together with the typical existence of two earners within a household. It is also close to $0.82$, the number reported by Chetty et al. (2011) in their meta-analysis of estimates for the Frisch elasticity using micro data. The value for $\chi$ is chosen[18] so that average hours worked equals $0.3$ (the associated average effective labor level, $N$, is $0.33$). To pin down the discount factor, $\beta$, we target a capital to output ratio of $2.72$, and the depreciation rate, $\delta$, is set to match an investment to output ratio of $27$ percent[19].

The aggregate technology is given by a Cobb-Douglas production function $Y = K^{\alpha}N^{1-\alpha}$ with capital share equal to $\alpha$, in the initial stationary equilibrium output is equal to $0.608$. The capital share parameter, $\alpha$, is set to its empirical counterpart of $0.38$.

## 3.2 Borrowing Constraints

We discipline the borrowing constraint $\underline{a}$ using the percentage of households in debt (negative net worth). We target $18.6$ percent following the findings of Wolff (2011) based on the $2007$ SCF.

[17] Marcet et al. (2007) investigate the role of wealth effects on the differences in allocation between complete and incomplete markets and conclude that they can be relevant under certain calibrations.

[18] It is understood that in any general equilibrium model all parameters affect all equilibrium objects. For the presentation purposes, we associate a parameter with the variable it affects quantitatively most.

[19] Capital is defined as nonresidential and residential private fixed assets and purchases of consumer durables. Investment is defined in a consistent way.

## 3.3 Fiscal policy

In order to set the tax rates in the initial stationary equilibrium we use the effective average tax rates computed by Trabandt and Uhlig (2011) from 1995 to 2007 and average them. The lump-sum transfers to output ratio is set to 8 percent and we discipline the government expenditure by imposing a debt to output ratio of 63 percent also following Trabandt and Uhlig (2011). The latter is close to the numbers used in the literature (e.g. Aiyagari and McGrattan (1998), Domeij and Heathcote (2004) or Winter and Roehrs (2016)). The calibrated value implies a government expenditure to output ratio of 15 percent, the data counterpart for the relevant period is approximately 18 percent. Further, we also approximate well the actual income tax schedule as can be seen in Figure 1.

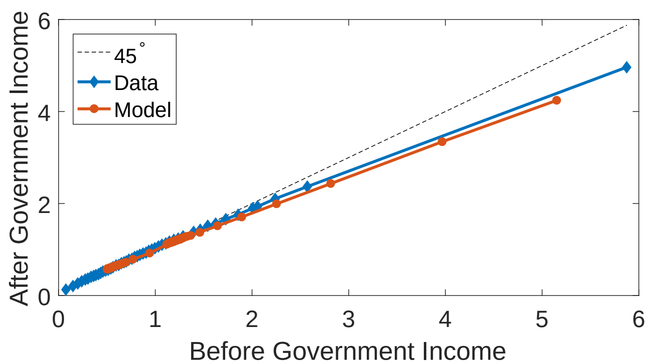



Figure 1: Income tax schedule

Notes: The data was generously supplied by Heathcote et al. (2014) who used PSID and the TAXSIM program to compute it. The axis units are income relative to the mean.

## 3.4 Labor income process

The individual labor productivity levels $e$ and transition probabilities in matrix $\Gamma$ are chosen to match the US wealth distribution, statistical properties of the estimated labor income process and the correlation between wealth and labor income. There are three levels of labor productivity in our model. Since we normalize the average productivity to one we are left with two degrees of freedom. The transition matrix is $3 \times 3$. The fact that it is a probability matrix implies its rows add up to one, therefore we are left with an additional six degrees of freedom. Thus, we end up with eight parameters to choose.

It is common to use the Tauchen method when calibrating the Markov process for productivities. This method imposes symmetry of the Markov matrix which further reduces the number of free parameters. Following Castañeda et al. (2003) we do not impose symmetry which allows us to target at the same time statistics from the labor income process and the individual wealth distribution.

To match the wealth distribution we target shares of wealth owned by the first, fourth and fifth quintile, the share of wealth owned by individuals in the top 5 percent and the Gini index. The targets are taken from the 2007 Survey of Consumer Finances[20]. We also target properties of individual labor income estimated as the AR(1) process, namely its autocorrelation and its standard deviation[21]. According to Domeij and Heathcote (2004), existing studies estimate the first order autocorrelation of (log) labor income to lie between 0.88 and 0.96 and the standard deviation (of the innovation term in the continuous representation) of 0.12 and 0.25. We calibrate the productivity process so that the Markov matrix and vector $e$ imply an autocorrelation of (log) labor income of 0.9 and a standard deviation of the innovation of 0.2[22] (in Section 7 we provide robustness results with respect to these choices). Finally, we target the correlation between wealth and labor income which is 0.29 in the 2007 SCF data. This way we discipline to some extent the labor income distribution using the wealth distribution that we match accurately. The resulting productivity vector, transition matrix and stationary distribution of productivities, $\lambda_e^*$, are

$$e = \begin{bmatrix} 0.791 \\ 1.266 \\ 4.938 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} .956 & .043 & .001 \\ .071 & .929 & .000 \\ .012 & .051 & .937 \end{bmatrix}, \quad \text{and} \quad \lambda_e^* = \begin{bmatrix} .616 \\ .377 \\ .007 \end{bmatrix}.$$

### 3.5 Model performance

Table 2 presents statistics about the wealth and labor income distributions. We target five of the wealth distribution statistics, so it is not surprising that we match that distribution quite well. Table 3 presents another crucial dimension along which our model is consistent with the data: income sources over the quintiles of wealth. The composition of income, specially of the consumption-poor agents, plays an important role in the determination of the optimal fiscal policy. The fraction of uncertain labor income determines the strength of the insurance motive and the fraction of the unequal asset income affects the redistributive motive. Our calibration delivers, without targeting, a good approximation of the income composition. Finally, we also match the consumption Gini which remained fairly constant around 0.27 in the period from 1995 to 2003 (see Krueger and Perri (2006)).

[20] For a general overview of this data see Díaz-Giménez et al. (2011).

[21] Including transitory shocks would allow a better match to the labor income process. However, these types of shocks can, for the most part, be privately insured against (see Guvenen and Smith (2013)) so we chose to abstract from them to keep the model parsimonious.

[22] We follow Nakajima (2012) in choosing these targets. The targets are associated with labor income, $wen$, which includes the endogenous variables $w$ and $n$. Therefore, to calibrate the parameters governing the individual productivity process, the model must be solved repeatedly until the targets are satisfied.

Table 2: Distribution of wealth

| | Bottom (%) | Quintiles | | | | | Top (%) | Gini |
|---|---|---|---|---|---|---|---|---|
| | 0-5 | 1st | 2nd | 3rd | 4th | 5th | 95-100 | |
| Data | -0.1 | -0.2 | 1.1 | 4.5 | 11.2 | 83.4 | 60.3 | 0.82 |
| Model | -0.1 | -0.2 | 1.5 | 5.1 | 10.2 | 83.4 | 60.8 | 0.81 |

Notes: Data come from the 2007 Survey of the Consumer Finances.

Table 3: Income sources by quintiles of wealth

| Quintile | Model | | | Data | | |
|---|---|---|---|---|---|---|
| | Labor | Asset | Transfer | Labor | Asset | Transfer |
| 1st | 83.8 | -0.3 | 16.5 | 81.9 | 2.1 | 16.1 |
| 2nd | 85.4 | 1.6 | 13.1 | 82.8 | 4.8 | 12.2 |
| 3rd | 84.1 | 4.7 | 11.2 | 80.0 | 7.3 | 12.6 |
| 4th | 81.4 | 8.6 | 10.0 | 77.6 | 10.3 | 12.2 |
| 5th | 58.7 | 36.1 | 5.1 | 51.8 | 40.0 | 8.2 |

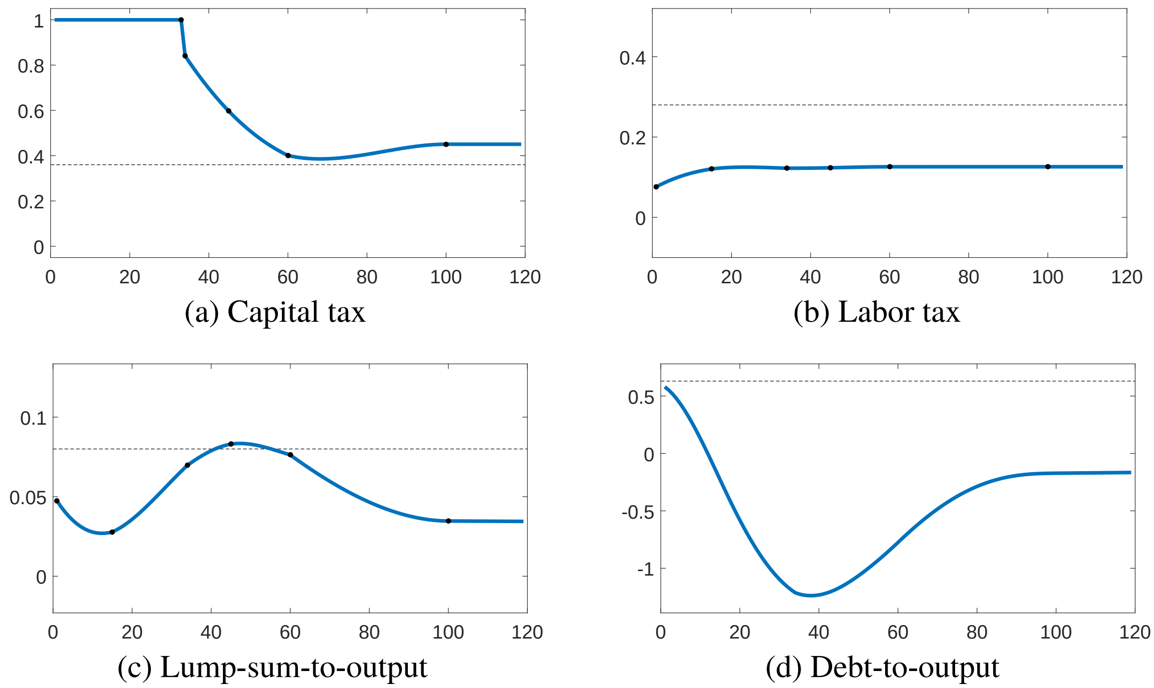
Notes: Table summarizes the pre-tax total income decomposition. We define the asset income as the sum of income from capital and business. Data come from the 2007 Survey of the Consumer Finances, the numbers are based on a summary by Díaz-Giménez et al. (2011).

# 4 Main Results

The optimal paths for the fiscal policy instruments are illustrated in Figure 2. Capital taxes are be front-loaded hitting the upper bound for $33$ initial periods then decrease to $45$ percent in the long run. Labor income taxes are substantially reduced to less than half of its initial level, from $28$ percent to about $13$ percent in the long run. The ratio of lump-sum transfers to output decreases initially to about $3$ percent, then increases back to its initial level of $8$ percent before it starts converging to its final level of $3.5$ percent. The government accumulates assets in the initial periods of high capital taxes reaching a level of debt-to-output of about $-125$ percent, which then converges to a final level of $-15$ percent. Relative to keeping fiscal instruments at their initial levels, this leads to a welfare gain equivalent to a permanent $4.7$ percent increase in consumption.

## 4.1 Aggregates

The aggregates associated with the implementation of the optimal policy are shown in Figure 7. The capital level initially decreases by about $8$ percent in the first $13$ years, but then increases towards a final level $20$ percent higher than in the initial steady state. The increase might be




(a) Capital tax

(b) Labor tax

(c) Lump-sum-to-output

(d) Debt-to-output

Figure 2: Optimal Fiscal Policy: Benchmark

Notes: Dashed line: initial stationary equilibrium; Solid line: optimal transition; The black dots are the choice variables: the spline nodes and $t^*$, the point at which the capital tax leaves the upper bound.

surprising at a first glance given the higher capital taxes. To understand is, first notice that, even if capital income taxes were set to $100$ percent forever, there would still be precautionary motives for the agents with relatively high productivity to save; if they receive a negative shock they can then consume their savings. The precautionary motive to save is actually strengthened since the (after-tax) labor income increases. The decrease in government debt also contributes substantially to this increase - an effect we explain further below in Section 4.5.4. Moreover, the level of aggregate labor increases by about $15$ percent immediately after the policy change following the reduction in labor taxes, increasing the marginal productivity of capital.

The higher levels of capital and labor lead to higher levels of output and consumption, which increases by $15$ and $20$ percent respectively over the transition. The accompanying increase in average consumption and labor has ambiguous effects on the welfare of the average agent. Hence, we also plot in Figure 7f what we call the average consumption-labor composite, defined below in equation (4.1), which is the more relevant measure for welfare. On impact the labor-consumption composite increases by $13$ percent as the higher consumption levels (due to the initial reduction in savings) more than compensate for the higher supply of labor. It then decreases for some periods following the reduction in output and the increasing savings. In the long run it returns to a level about $13$ percent higher than in the initial steady state.

## 4.2 Long-Run Optimality Conditions

Aiyagari (1995) analyses the optimal long-run capital taxes in an environment similar to the one we are working with.[23] He argues that, since there is no aggregate uncertainty, the Ramsey planner's decision to move resources across time is risk-free and the associated Euler equation, in the long run, implies the modified golden rule (i.e. $\beta(1+f_K(K,N)) = 1$). On the other hand, agents face idiosyncratic shocks and the possibility of being borrowing-constrained in some future periods which leads to extra savings due to precautionary reasons. In order to implement the optimal level of capital in the long run it follows that the planner must set positive capital taxes. This logic also implies that the modified golden rule should hold in the long run; our numerical results imply exactly that. Figure 3 displays $\beta(1+f_K(K,N))$ for our benchmark results (solid line) and for an experiment, described in more detail in Section 4.6.2, in which we restrict the policy instruments to remain constant throughout the transition (dashed line). It becomes clear that the variations of taxes over time are crucial to approximate the long-run properties of the optimal tax system. Moreover, we view this as corroborating evidence for the accuracy of our numerical long-run results.

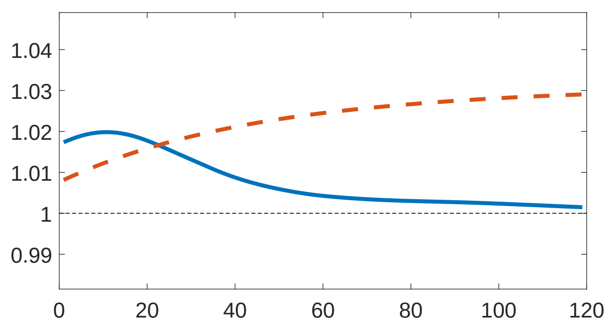



Figure 3: $\beta(1+f_K(K,N))$

Notes: Solid line: benchmark experiment; Dashed line: optimal transition with constant policy.

Recently Acikgoz (2015) and Hagedorn et al. (2016) have made advancements towards obtaining a better characterization of the long-run optimal tax system in environments very similar to ours. Both papers claim that the long-run optimal tax system is independent of initial conditions and of the transition towards it.[24] Moreover, they show that three optimality conditions must be satisfied (the modified golden rule and two additional ones) and propose an algorithm that allows for the computation of the optimal long-run tax system. We have applied this algorithm to our

[23]The home production assumption in Aiyagari (1995) is equivalent to our assumption that preferences are GHH. The differences are that in his environment the planner does not have lump-sum taxes as an instrument, but chooses the level of government expenditure every period (which enters separably in the agents' utility functions).

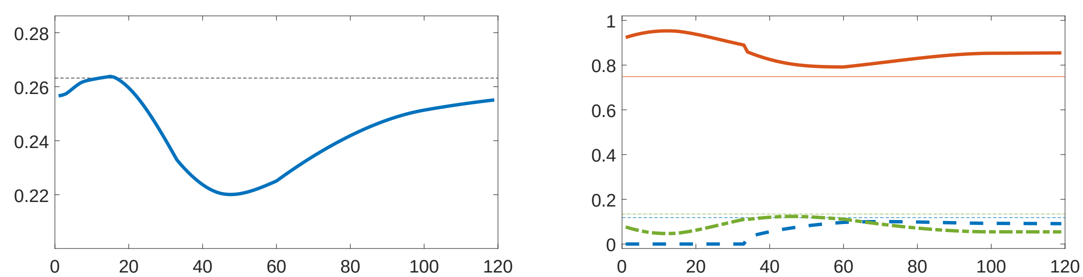
[24]At the time of writing, a formal proof is not available, though a convincing heuristic argument has been made for it, see Acikgoz (2015).

economy and have found very similar results[25]. We view this as further corroborating evidence of the accuracy of our (and their) results. It is reassuring that, even though long-run taxes only affect welfare in the far future[26], our algorithm is still able to accurately approximate long-run optimal taxes.

## 4.3 Distributional Effects

Movements in the levels do not provide a full picture of what results from the implementation of the optimal fiscal policy. It is also important to understand its effects on inequality and on the risk faced by the agents. Figure 4a plots the evolution of the Gini index for consumption[27]. Notice that, though it takes some time for the reduction to start, the consumption Gini is significantly reduced over the transition reaching a low about 16 percent lower than the initial level. As will become clear below, this reduction in inequality is behind most of the welfare gains associated with the optimal policy. Not surprisingly, such a change would be supported by most agents in the economy with the exception of the highly productive and, therefore, wealthier ones - see Table 4.

Figure 4b displays the evolution of the shares of labor, capital and transfer income out of total income. Importantly, notice that the share of labor income is significantly increased under the optimal policy. Since all the risk faced by agents in the SIM model is associated with their labor income, it turns out that they face more risk after the policy is implemented. This has an obvious negative effect on welfare which is, however, outweighed by the gains associated with the higher levels of consumption and the reduction in inequality it provides. The next sections will clarify some of these issues.




(a) Consumption Gini (b) Income shares

Figure 4: Inequality measures

Notes (a) and (b): Thin lines: initial stationary equilibrium; Thick lines: optimal transition. Notes (b): Solid lines: labor income share; Dash-dotted lines: transfer income share; Dashed lines: asset income share

[25]See the Online Appendix for details on how the conditions can be adapted to our environment exactly and for the results we obtain using them.

[26]The last time period we allow taxes to change is at $t = 100$, and $\sum_{t=100}^{\infty} \beta^t / \sum_{t=0}^{\infty} \beta^t \approx 0.03$.

[27]Since labor supply is proportional to productivity levels, the inequality of hours is unaffected by the policy, it is in fact determined exogenously. Hence, here we can focus on consumption inequality.

Table 4: Proportion in favor of reform

| $e = L$ | $e = M$ | $e = H$ | All |
|---|---|---|---|
| 99.5 | 98.3 | 4.2 | 99.4 |

### 4.4 Welfare decomposition

Here we present a result that will be particularly helpful for understanding the properties of the optimal fiscal policy. First, let $x_t$ be the individual consumption-labor composite (the term inside the utility function 3.1), that is

$$x_t \equiv c_t - \chi \frac{n_t^{1+\frac{1}{\kappa}}}{1+\frac{1}{\kappa}}, \tag{4.1}$$

and $X_t$ denote its aggregate level. The utilitarian welfare function can increase for three reasons. First, it will increase if the utility of the average agent, $U(\{X_t\})$, increases; we call this the *level effect*. Reductions in distortive taxes will achieve this goal by allocating resources more efficiently. This is the only relevant effect in a representative agent economy (without heterogeneity). Second, since agents are risk averse, it increases if the uncertainty about individual paths $\{x_t\}_{t=0}^{\infty}$ is reduced; we call this the *insurance effect*. By redistributing from the (ex-post) lucky to the (ex-post) unlucky, a tax reform reduces the uncertainty faced by the agents. Finally, it will increase if the inequality across the certainty equivalents of the individual paths $\{x_t\}_{t=0}^{\infty}$, for agents with different initial (asset/productivity) states, is reduced; we call this the *redistribution effect*. By redistributing from the rich (ex-ante lucky) to the poor (ex-ante unlucky), the tax reform reduces the inequality between agents. In Appendix B we give precise definitions for each of these effects and show how it is possible to measure them. Then, letting $\Delta$ be the average welfare gain, $\Delta_L$ the gains associated with the level effect, $\Delta_I$ with the insurance effect, and $\Delta_R$ with the redistribution effect, we prove the following proposition.

**Proposition 3** *If preferences are GHH as in* (3.1)*, then*

$$1 + \Delta = (1 + \Delta_L)(1 + \Delta_I)(1 + \Delta_R).$$

**Proof.** See Appendix B. ■

Hence, it is possible to decompose the average welfare gains into the components described above[28]. The results for this decomposition for our main results are in Table 5. Most of the

[28]The welfare gains described above are in terms of consumption-labor composite units. The decomposition does not hold exactly in terms of consumption units. To keep our results comparable with others, we report the average welfare gains in terms of consumption units and rescale the numbers for $\Delta_L$, $\Delta_I$, and $\Delta_R$ accordingly.

welfare gains implied by the implementation of the optimal fiscal policy come from the reduction in ex-ante inequality (redistribution effect). The also substantial welfare gains associated with the reduction in distortions (level effect) is almost exactly offset by welfare losses due to the increase in uncertainty (insurance effect). The labor income process hard-wires a strong savings motive for the high income agents, since they have a low probability of reaching such a state and a high probability of leaving it. It follows that the savings decision is less elastic than the labor supply decision, and replacing labor taxes with capital taxes reduces overall distortions.

Table 5: Welfare decomposition

| Average welfare gain $\Delta$ | Level effect $\Delta_L$ | Insurance effect $\Delta_I$ | Redistribution effect $\Delta_R$ |
|---|---|---|---|
| 4.7 | 3.1 | -4.5 | 6.4 |

## 4.5 Fixed instruments

In order to understand the role played by each instrument in the optimal fiscal policy, we ran experiments in which we hold each of them fixed and optimize only with respect to the others. Table 6 displays the welfare decomposition for each of these experiments.[29]

### 4.5.1 Capital taxes

It is clear from the welfare decomposition in Table 6 that the path of capital taxes plays a crucial role in the redistributional gains associated with the unrestricted optimal policy. Restricting capital taxes to their initial level brings the redistribution effect from $6.4$ percent to $-0.4$ percent. In line with the result in Proposition 2, the increase in capital taxes especially in the initial years leads to a strong redistribution effect as the proportion of unequal asset income is reduced (actually brought to zero in the first $33$ years). Relative to the optimal policy, the restriction on capital taxes also leads to higher labor taxes (which explains the better insurance effect) and a lower accumulation of assets by the government.

### 4.5.2 Labor taxes

Fixing labor taxes at their initial level is particularly detrimental to the level effect. In the optimal policy labor taxes are reduced substantially and the labor supply distortions reduced accordingly, which drives up the capital stock by increasing it's marginal product. The distributional gains are

[29]The corresponding figures can be found in the Online Appendix.

virtually unaffected whereas the insurance effect is improved, which is consistent with the result in Proposition 1 since the restriction implies higher labor taxes. The fact that the insurance effect is still negative might be surprising though. What is behind this effect is the role played by the accumulation of assets by the government which we explain bellow.

### 4.5.3 Lump-sum transfers

Restricting lump-sum transfers to its initial level does not affect the results as much as the other restrictions; the average welfare gains are reduced from $4.7$ percent to $4.2$ percent. Most of the losses come from the reduction in the level effect. The restriction leads to a higher overall level of transfers and, therefore, higher labor taxes relative to the unrestricted optimal policy whereas capital taxes are virtually unaffected. This leads to an overall higher level of distortions which explains the lower level effect.

Table 6: Welfare decomposition: Fixed instruments

| | $\Delta$ | $\Delta_L$ | $\Delta_I$ | $\Delta_R$ |
|---|---|---|---|---|
| Fixed capital taxes | 1.0 | 3.6 | -2.2 | -0.4 |
| Fixed labor taxes | 3.2 | -0.9 | -2.2 | 6.5 |
| Fixed lump-sum | 4.2 | 0.9 | -3.2 | 6.7 |
| Fixed debt | 3.8 | 2.6 | -2.9 | 4.2 |
| **Benchmark** | 4.7 | 3.1 | -4.5 | 6.4 |

### 4.5.4 Government debt

In the absence of borrowing constraints an increase in government debt is innocuous, in response agents simply adjust their savings one-to-one and the Ricardian equivalence holds. In the SIM model, however, agents face a borrowing constraint (which is binding for some of them). The Ricardian equivalence breaks down and in response to an increase in government debt aggregate savings increase by less than one-to-one. Since the asset market must clear (i.e. $A_t = K_t + B_t$), it follows that capital must decrease as a result. Hence, increases in government debt crowd out capital while decreases crowd in capital[30].

In order to understand why the government accumulates assets in the optimal policy it is important to look at its effect on equilibrium prices[31]. A lower amount of government debt leads

[30]See Aiyagari and McGrattan (1998) and Winter and Roehrs (2016) for an extensive discussion of this issue.

[31]The fact that the government accumulates assets does not imply that it becomes the owner of part of the capital stock. Agents own the capital, but on average owe the government (in the form of IOU contracts) more than the value of their capital holdings.

to a higher level of capital which reduces interest rates and increases wages. Hence, besides the potential positive level effect associated with the higher levels of capital such a policy also affects the insurance and redistribution effects. It effectively reduces the proportion of the agents' income associated with the unequal asset income and increases the proportion associated with uncertain labor income. The result is a positive redistribution effect and a negative insurance effect. Thus, when government debt-to-output is held fixed the redistributional gains are reduced from $6.4$ percent to $4.2$ percent while the insurance loss is reduced from $-4.5$ percent to $-2.9$ percent. This also clarifies why the planner chooses to accumulate assets when the instrument is not restricted: the welfare gains associated with the resulting redistribution outweigh the losses from the increased uncertainty.

## 4.6 Transitory effects

In this section we first compute the optimal fiscal policy ignoring transitory welfare effects. A comparison with our benchmark results allows us to measure the importance of accounting for these effects. If the difference was small this would be a validation of experiments of this kind performed in the literature. It turns out, however, that the results are remarkably different. A better option, is to solve for the optimal policy with constant instruments accounting for transitory welfare effects. The welfare loss associated with holding the instruments constant, however, is still significant. The results are summarized in Tables 7 and 8.

Table 7: Final Stationary Equilibrium: transitory effects

| | $\tau^h$ | $\tau^k$ | $T/Y$ | $B/Y$ | $K/Y$ | $N$ | $r$ | $w$ |
|---|---|---|---|---|---|---|---|---|
| Initial equilibrium | 28.0 | 36.0 | 8.0 | 63.0 | 2.71 | 0.33 | 4.1 | 1.14 |
| Stat. equil. | 18.0 | - | 3.73 | -326.0 | 3.99 | 0.43 | -0.4 | 1.45 |
| Stat. equil. fixed debt | 4.7 | -5.2 | -5.4 | 43.2 | 3.20 | 0.43 | 1.9 | 1.26 |
| Constant policy | 7.7 | 73.7 | 3.5 | 53.4 | 2.21 | 0.36 | 7.2 | 1.01 |
| **Benchmark** | 12.6 | 45.1 | 3.4 | -15.1 | 2.82 | 0.39 | 3.5 | 1.17 |

Notes: The values of $\tau^h$, $\tau^k$, $T/Y$, $B/Y$, and $r$ are in percentage points.

### 4.6.1 Maximizing steady state welfare

Here the the planner chooses stationary levels of all four fiscal policy instruments to maximize welfare in the final steady state. In particular, the planner can choose any level of government debt without incurring in the transitional costs associated with it. It chooses a debt-to-output ratio of $-326$ percent. At this level the amount of capital that is crowded in is close to the golden

Table 8: Welfare decomposition: transitory effects

| | $\Delta$ | $\Delta_L$ | $\Delta_I$ | $\Delta_R$ |
|---|---|---|---|---|
| Stat. equil. | 24.8 | 28.9 | -12.0 | 9.9 |
| Stat. equil. fixed debt | 9.2 | 26.7 | -9.5 | -4.7 |
| Constant policy | 3.2 | 2.6 | -3.1 | 3.8 |
| **Benchmark** | 4.7 | 3.1 | -4.5 | 6.4 |

rule level, that is, such that interest rates (net of depreciation) equal to zero. Thus, taxing capital income in this scenario has no relevant effect and we actually find multiple solutions with different levels of capital taxes which is why we do not display that number in Table 7. The average welfare gains associated with this policy are of $24.8$ percent, that is, agents would be willing to pay this percentage of their consumption in order to be born in the stationary equilibrium of an economy that has this policy instead of the initial stationary equilibrium. However, these welfare gains ignore the transitory effects, it is as if the economy jumped immediately to a new steady state in which the government has a large amount of assets without incurring in the costs associated with accumulating it.

A more reasonable experiment, which is closer to the one studied by Conesa et al. (2009), is to restrict the level of debt to remain at its initial level. When this is the case, the planner reduces labor taxes and capital taxes substantially obtaining most of the necessary revenue via lump-sum taxes. This has detrimental insurance and redistribution effects, but the associated level effect more than makes up for it. The policy leads to a welfare gain of $9.2$ percent relative to the initial steady state when transitory effects are ignored. However, once transitory effects are considered, implementing this policy leads to a welfare *loss* of $6.4$ percent. Hence, ignoring transitory effects can be severely misleading. Importantly, the transitory distributional effects of the policy and the costs associated with the accumulation of capital (or assets by the government) are ignored.

### 4.6.2 Transition with constant policy

Here we consider the problem of finding the *constant* optimal fiscal policy that maximizes the same welfare function we use in our benchmark experiment, in which transitory effects are accounted for.[32] The level of capital taxes is close to average between the upper bound of $100$ percent and the final capital tax in the benchmark experiment. Labor taxes are reduced from a long-run level of $12.6$ percent to $7.7$ percent and lump-sum transfers converge much faster to the final level of $3.5$ percent. The main difference in the fiscal policy instruments is the fact that with a con-

[32] We present figures comparing these results with the benchmark ones in the Online Appendix.

stant policy the government is not able to accumulate assets via higher initial capital taxes. The debt-to-output ratio remains close to the initial level, even though we do not impose any restrictions on it in this experiment. As a result of the higher long-run capital tax and relatively higher debt-to-output ratio, capital decreases by about $20$ percent in the long run whereas it *increases* by approximately the same amount in the benchmark experiment. The associated higher interest rates and lower wages lead to the reduction in the redistributional gains and reduces the insurance losses associated with the lower labor tax. This policy leads to an average welfare gain of $3.2$ percent whereas the time varying policy increases welfare by $4.7$ percent. That is, the restriction to constant policies leads a welfare loss of $1.5$ percent.

# 5 Complete Market Economies

To our knowledge, this paper is the first to solve the Ramsey problem in the SIM environment. To highlight the role of the market incompleteness for the optimal policy and relate our findings to other results in the literature, we provide a build up to our benchmark result. First, we start from the representative agent economy (Economy 1) and introduce heterogeneity only in initial assets (Economy 2), heterogeneity only in individual productivity levels (constant and certain) (Economy 3), and heterogeneity both in initial assets and in individual productivity levels (Economy 4). Introducing idiosyncratic productivity shocks and borrowing constraints brings us back to the SIM model. At each step, we analyze the optimal fiscal policy identifying the effect of each feature.

In what follows we examine the optimal fiscal policy in Economies 1-4. Their formal environments can be quickly described by starting from the SIM environment delineated above. Economy 4 is the SIM economy with transition matrix, $\Gamma$, set to the identity matrix. and borrowing constraints replaced by no-Ponzi conditions. Then, we obtain Economy 3 by setting initial asset levels to its average, Economy 2 by setting the productivity levels to its average, $e = 1$, and Economy 1 by equalizing both initial assets and levels of productivity. Figure 5 contains the numerical results obtained using the same method used for the benchmark results together with some of the analytical equations derived bellow.

## 5.1 Economy 1: representative agent

To avoid a trivial solution, the usual Ramsey problem in the representative agent economy does not consider lump-sum transfers to be an available instrument. Since in this paper we do, the solution is, in fact, very simple. It is optimal to obtain all revenue via lump-sum taxes and set capital and labor income taxes so as not to distort any of the agent's decisions. This amounts to $\tau_t^k = 0$ and $\tau_t^n = -\tau^c$ for all $t \geq 1$. Since consumption taxes are exogenously set to a constant level, zero capital taxes leaves savings decisions undistorted and labor taxes equal to minus the consumption tax ensures labor supply decisions are not distorted as well. In this setup the Ricardian equivalence holds, so that the optimal paths for lump-sum taxes and debt are indeterminate: there is no lesson to be learned from this model about the timing of lump-sum taxes or the path of government debt. This will also be the case in Economies 2, 3 and 4 and is why we do not discuss or plot them.

### 5.2 Economy 2: heterogeneity in initial assets

Introducing heterogeneity in the initial level of assets we can diagnose the effect of this particular feature on the Ramsey policies by comparing it to the representative agent ones. We extend the procedure introduced by Werning (2007)[33] to characterize the optimal policies for this and the next two economies. For the economy with heterogeneity in asset we obtain the following proposition.

**Proposition 4** *There exists a finite integer* $t^* \geq 1$ *such that the optimal tax system is given by* $\tau_t^k = 1$ *for* $1 \leq t < t^*$ *and* $\tau_t^k = 0$ *for all* $t > t^*$*; and* $\tau_t^n = -\tau^c$ *for all* $t \geq 1$.

**Proof.** See Appendix C.1. ■

The results in this and the next two propositions are valid for any set of welfare weights.[34] Hence, we effectively characterize the set of Pareto efficient policies. In this Proposition, in particular, a change in the welfare weights would only change $t^*$, leaving unchanged the long run optimal levels of capital and labor income taxes. In a similar setting Laczo et al. (2015) obtain analogous results. In Section 6 we show that the long-run taxes in the benchmark results are also robust to some changes in the welfare weights.

Once again, there is no reason to distort labor decisions since labor income is certain and the same for all agents. However, the path of capital taxes differs from the representative agent ones. Proposition 2 provides a rationale for taxing capital in this case; since agents have different initial asset levels, capital taxes can be used to provide redistribution. This fact together with the fact that capital taxes are zero in the long run determine the optimal path for capital taxes[35]. Capital taxes are positive and front-loaded, hitting the upper bound in the initial periods and subsequently being driven to zero. The extra revenue obtained via capital taxation is redistributed via lump-sum transfers (or a reduction in lump-sum taxes relative to the representative agent level). It is important to reemphasize that since lump-sum transfers are an unrestricted instrument, there is no reason to tax capital in the initial periods other than for redistributive motives.

[33] Werning (2007) solves for separable and balance growth path utility functions. Besides solving for GHH preferences we also impose the upper bound on capital income taxes and remove the possibility of time zero taxation to keep the results comparable with the benchmark ones.

[34] The associated numerical results do assume a utilitarian welfare function.

[35] Straub and Werning (2014) show that optimal long-run capital taxes can be positive in environments similar to this one. The reason why their logic does not apply here is the fact that the planner has lump-sum taxes as an available instrument which removes the need to obtain revenue via distortive instruments. In the Online Appendix we include a more detailed discussion of this issue.

In order to have a sense of the magnitudes of $t^*$ and the increase in lump-sum transfers, we apply the same procedure to the one we used to solve for the optimal tax system in the benchmark economy. All we need to do is choose the initial distribution of assets. The stationary distribution of assets in this economy is indeterminate[36], hence, we can choose any one we want. To keep the results comparable we choose the initial stationary distribution from the benchmark experiment[37].

### 5.3 Economy 3: heterogeneity in productivity levels

It turns out that the Ramsey policies for this economy are a bit more complex. Let $\Phi$, $\Psi$, and $\Omega^n$ be constants (defined in Appendix C) and define

$$\Theta_t \equiv \frac{C_t}{\Omega^n \chi_{\frac{\kappa}{1+\kappa}} N_t^{\frac{1+\kappa}{\kappa}}} - 1.$$

The following proposition can be established.

**Proposition 5** *Assuming capital taxes are bounded only by the positivity of gross interest rates, the optimal labor tax, $\tau_t^n$, can be written as a function of $\Theta_t$ given by*

$$\tau_t^n(\Theta_t) = \frac{(1+\tau^c)\Psi\Theta_t}{\Phi\Theta_t + \Psi(\sigma + \Theta_t)} - \tau^c, \quad \textit{for } t \geq 1, \tag{5.1}$$

*with sensitivity*

$$\Theta_t \frac{d\tau_t^n(\Theta_t)}{d\Theta_t} = \frac{\sigma(\tau_t^n(\Theta_t) + \tau^c)^2}{(1+\tau^c)\Theta_t}. \tag{5.2}$$

*It is optimal to set the capital-income tax rate according to*

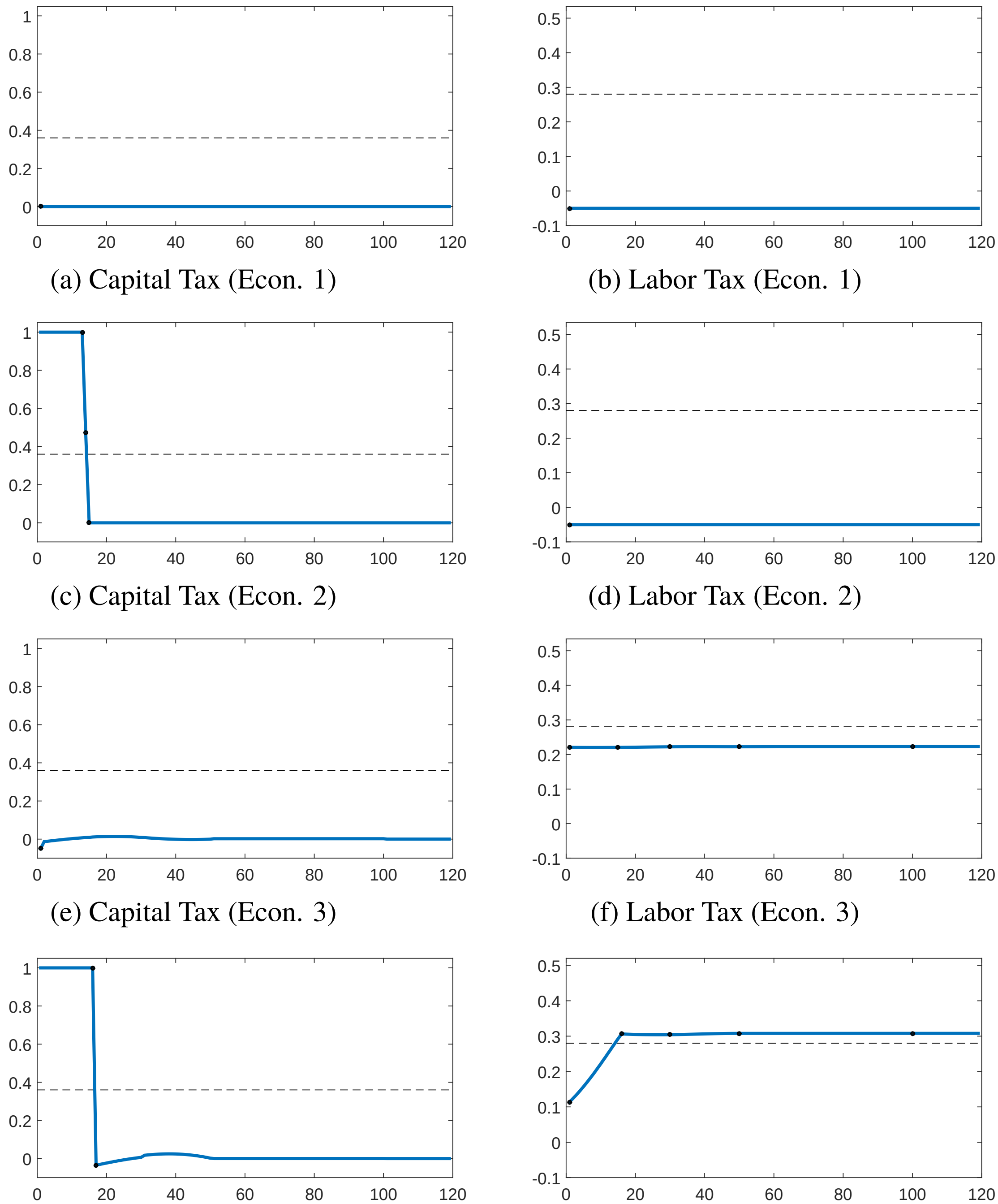
$$\frac{1 + (1-\tau_{t+1}^k) r_{t+1}}{1 + r_{t+1}} = \frac{\tau_t^n + \tau^c}{\tau_{t+1}^n + \tau^c} \frac{1 - \tau_{t+1}^n}{1 - \tau_t^n}, \quad \textit{for } t \geq 1. \tag{5.3}$$

**Proof.** See Appendix C.2. ■

Since labor income is unequal, there is a redistributive reason to tax it. Optimal labor taxes are not constant over time since they depend on $\Theta_t$. If they were constant, however, equation (5.3) would imply $\tau_t^k = 0$ for all $t \geq 2$. Thus, capital taxes will fluctuate around zero to the extent that labor taxes vary over time. We disregard the upper bound on capital taxes, $\tau_{t+1}^k \leq 1$, because it would complicate the result even further and in a non-interesting way. It could be that the bound

[36] For the preferences chosen above, consumption is linear on the individual asset level, and labor supply is independent of it. It follows that the equilibrium levels of aggregates are independent of the asset distribution and equal to the representative agent ones (see Chatterjee (1994)). In a steady state, $\beta\left(1 + \left(1 - \tau^k\right) r\right) = 1$ and, therefore, every agent will keep its asset level constant.

[37] In fact, a rescaling of it since the steady state aggregate level of assets is different when there is no idiosyncratic risk (since there is no precautionary motive for savings).



(a) Capital Tax (Econ. 1) (b) Labor Tax (Econ. 1)

(c) Capital Tax (Econ. 2) (d) Labor Tax (Econ. 2)

(e) Capital Tax (Econ. 3) (f) Labor Tax (Econ. 3)

(g) Capital Tax (Econ. 4) (h) Labor Tax (Econ. 4)

Figure 5: Optimal Taxes: Complete Market Economies

Notes: Dashed line: initial taxes; Solid line: optimal taxes.

is violated if the variation of $\Theta_t$ between $t$ and $t+1$ is large enough. However, as discussed below, quantitatively this is unlikely.

To obtain a numerical solution we set the productivity levels to the ones in the benchmark economy and apply the same procedure. To have a sense of the magnitude of the sensitivity of $\tau_t^n$ to $\Theta_t$ we plug the initial stationary equilibrium numbers ($\tau^n = 0.28$, $\tau^c = 0.05$, $\sigma = 2$, and $\Theta \approx 3$) into equation (5.2). This implies a sensitivity of $0.07$, i.e. a $1$ percent increase in $\Theta_t$ changes the tax rate by $0.06$ of a percentage point, from $0.28$ to $0.2798$. [38] Notice that the volatility of $\Theta_t$ over time is unsubstantial. It follows that the optimal labor taxes are virtually constant and capital taxes virtually zero.

In any case, the fact that capital is taxed at all seems to be inconsistent with the logic put forward so far. It is not. When labor taxes vary over time they distort the savings decision, capital taxes are then set to "undo" this distortion. The analogous is not the case in Economy 2 because of the absence of income effects on labor supply; distortions of the savings decision do not affect the labor supply.

### 5.4 Economy 4: heterogeneity in initial assets and productivity levels

The result for this economy is a combination of the last two.

**Proposition 6** *There exists a finite integer $t^* \geq 1$ such that the optimal tax system is given by $\tau_t^k = 1$ for $1 \leq t < t^*$, $\tau_t^k$ follows equation (5.3) for $t > t^*$; $\tau_t^n$ evolves according to equation (5.3) for $1 \leq t < t^*$; and $\tau_t^n$ is determined by equation (5.1) for all $t \geq t^*$.*

**Proof.** See Appendix C.3. ■

Optimal capital taxes are very similar to Economy 2 and for the same reasons. Labor taxes are determined by the same equation as in Economy 3 for $t \geq t^*$. In initial period, $1 \leq t < t^*$, while capital taxes are at the upper bound, $R_t = 1 < R_t^*$ and, therefore, equation (5.3) implies that labor taxes should be increasing. Lump-sum transfers are higher than the in Economies 2 and 3 since they are used to redistribute the capital *and* labor tax revenue.

## 6 Controlling the degree of inequality aversion

Figure 6 shows that the solution with $4$ nodes ($t^*$,$\tau_{t^*+1}^k$,$\tau_1^n$, and $T_1$) produces a reasonable approximation for the benchmark solution, at least with respect to its basic features, it leads to welfare gains of $4.65$ percent relative to $4.74$ percent in the benchmark results. In this section, we use this approximation to explore the effects of changing the planner's degree of inequality aversion.

[38] We can also calculate the path of $\Theta_t$, which we displayed in a figure in the Online Appendix.

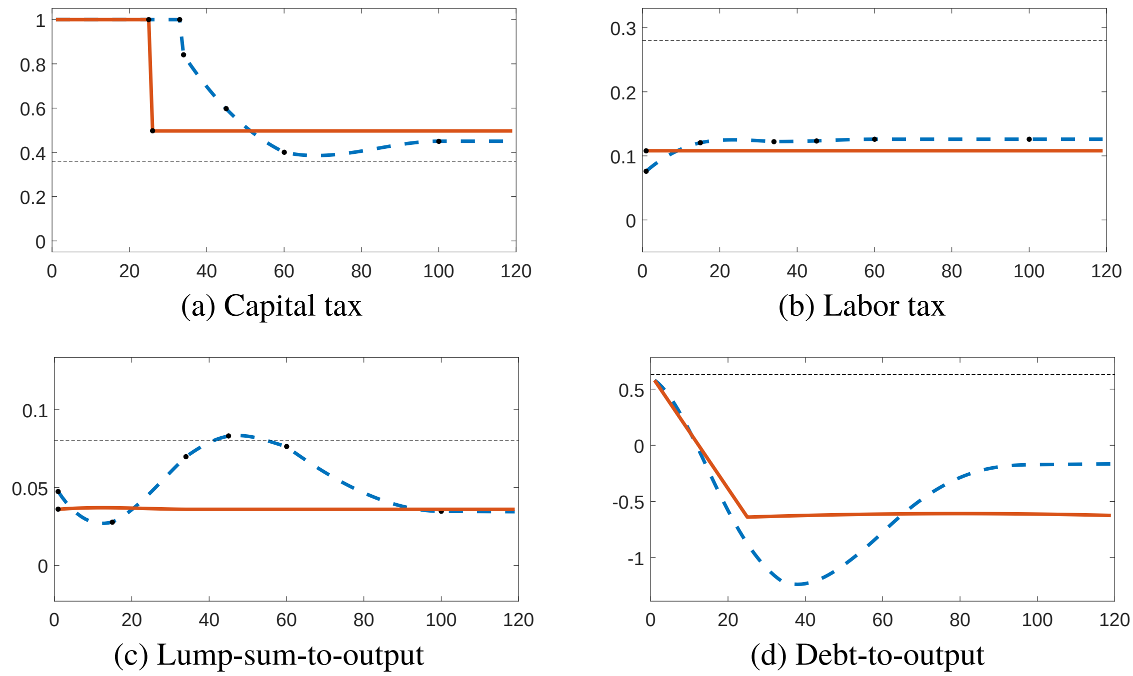



Figure 6: Optimal Fiscal Policy with 4 nodes

Notes: Dashed thin line: initial stationary equilibrium; Dashed thick line: optimal transition with 17 nodes (benchmark); Solid line: optimal transition with 4 nodes.

All the results presented so far used the same social welfare function: the utilitarian one, which places equal Pareto weights on each agent. This implies a particular social preference with respect to the equality versus efficiency trade-off. Here we consider different welfare functions that rationalize different preferences about this trade-off. With this in mind we propose the following function

$$W^{\hat{\sigma}} = \left( \int \bar{x} \left( a_0, e_0 \right)^{1-\hat{\sigma}} d\lambda_0 \right)^{\frac{1}{1-\hat{\sigma}}},$$

where $\lambda_0$ is the initial distribution of individual states $(a_0, e_0)$, $\bar{x}$ denotes the individual certainty equivalents of labor-consumption composite (given a particular initial state $(a_0, e_0)$), and, following Benabou (2002), we call $\hat{\sigma}$ the planner's degree of inequality aversion. First notice that if $\hat{\sigma} = \sigma$ (the agents' degree of risk aversion), maximizing $W^{\sigma}$ is equivalent to maximizing the utilitarian welfare function [39]. If $\hat{\sigma} = 0$, then maximizing $W^{0}$ is equivalent to maximizing $(1 + \Delta_L)(1 + \Delta_I)$, that is, the planner has no redistributive concerns and focuses instead in the reduction of distortions and the provision of insurance[40]. Finally, as $\hat{\sigma} \to \infty$ the welfare func-

[39]Notice that $\left( \int \bar{x} \left( a_0, e_0 \right)^{1-\sigma} d\lambda_0 \right)^{\frac{1}{1-\sigma}}$ is a monotonic transformation of $\int \frac{\bar{x}(a_0,e_0)}{1-\sigma}^{1-\sigma} d\lambda_0$, which is equivalent to the utilitarian welfare function.

[40]This result can be established following a similar procedure to the one used in proof of Proposition 3. The Online

tion approaches $W^{\infty} = \min(\bar{x}(a_0, e_0))$. Hence, by choosing different levels for $\hat{\sigma}$ we can place different weights on the equality versus efficiency trade-off, from the extreme of completely ignoring equality ($\hat{\sigma} = 0$), passing through the utilitarian welfare function ($\hat{\sigma} = \sigma$), and in the limit reaching the Rawlsian welfare function ($\hat{\sigma} \to \infty$). Table 9 displays the results for different levels of $\hat{\sigma}$.

Table 9: Controlling the degree of inequality aversion

| | $t^*$ | $\tau^k$ | $\tau^n$ | $T/Y$ | $B/Y$ | $\Delta$ | $\Delta_L$ | $\Delta_I$ | $\Delta_R$ |
|---|---|---|---|---|---|---|---|---|---|
| Degree of Inequality Aversion. Benchmark: $\hat{\sigma} = 2$ | | | | | | | | | |
| $\hat{\sigma} = 0.0$ | 0 | 34.6 | 12.2 | 0.0 | 79.8 | 0.59 | 5.49 | -2.87 | -1.83 |
| $\hat{\sigma} = 1.0$ | 19 | 49.9 | 10.1 | 2.9 | -36.4 | 4.57 | 3.86 | -3.94 | 4.80 |
| $\hat{\sigma} = 2.0^*$ | 26 | 49.7 | 10.8 | 3.6 | -62.9 | 4.65 | 3.09 | -3.94 | 5.66 |
| $\hat{\sigma} = 3.0$ | 29 | 49.7 | 10.4 | 3.5 | -76.8 | 4.64 | 3.00 | -4.09 | 5.93 |
| $\hat{\sigma} = 4.0$ | 30 | 48.9 | 11.5 | 4.1 | -76.0 | 4.61 | 2.61 | -3.86 | 6.05 |
| $\hat{\sigma} = 5.0$ | 32 | 49.2 | 11.3 | 4.0 | -84.2 | 4.59 | 2.53 | -3.95 | 6.21 |

Notes: When $\hat{\sigma} = 2 = \sigma$ the welfare function is utilitarian, this is the solution plotted in Figure 6. The values for $T/Y$ and $B/Y$ are the ones from the final steady state. For the welfare decomposition we use the utilitarian welfare function for comparability.

When $\hat{\sigma} = 0$ the planner has no redistributive motive and, accordingly, $t^* = 0$ which is consistent with the results displayed above, in particular in Section 5. The benchmark result that capital taxes should be held fixed at the upper bound for the initial periods is inherently linked to the redistributive motive of the planner. It follows that higher $\hat{\sigma}$ imply higher $t^*$'s (lower lump-sum-to-output ratios and higher debt-to-output ratios). Otherwise, overall, specially for $\hat{\sigma} \geq 1$, the results do not change significantly with changes in $\hat{\sigma}$. In particular, the final levels of capital and labor taxes are remarkably similar.

# 7 Robustness

In this section we use the same 4-node approximation used in the previous one to evaluate the robustness of the results with respect to the labor income process and key elasticities.

## 7.1 Labor income process

The labor income process (summarized by the Markov matrix, $\Gamma$, and the vector of productivity levels, $e$) is a key determinant of the amount of uncertainty and inequality faced by agents in

Appendix contains the proof.

Table 10: Robustness

| | $t^*$ | $\tau^k$ | $\tau^n$ | $T/Y$ | $B/Y$ | $\Delta$ | $\Delta_L$ | $\Delta_I$ | $\Delta_R$ |
|---|---|---|---|---|---|---|---|---|---|
| Labor Income Process. Benchmark: $\rho = 0.9$, $\sigma_\varepsilon = 0.2$ | | | | | | | | | |
| $\rho = 0.85$ | 24 | 34.8 | 4.8 | 0.0 | -100.1 | 5.43 | 4.97 | -3.86 | 4.47 |
| $\rho = 0.95$ | 21 | 42.8 | 11.5 | 3.7 | -49.5 | 3.90 | 3.54 | -3.29 | 3.76 |
| $\sigma_\varepsilon = 0.15$ | 28 | 28.1 | 4.9 | 0.1 | -126.3 | 5.64 | 4.67 | -4.16 | 5.32 |
| $\sigma_\varepsilon = 0.25$ | 34 | 57.8 | 11.6 | 4.7 | -75.9 | 4.52 | 2.55 | -4.32 | 6.52 |
| Degree of Relative Risk Aversion and Frisch Elasticity. Benchmark: $\sigma = 2$, $\kappa = 0.72$ | | | | | | | | | |
| $\sigma = 1.0$ | 12 | 25.0 | 9.9 | 0.3 | -21.7 | 3.75 | 5.82 | -3.44 | 1.54 |
| $\sigma = 3.0$ | 50 | 74.4 | 10.0 | 5.1 | -93.2 | 6.75 | 1.90 | -3.80 | 8.90 |
| $\kappa = 0.5$ | 24 | 49.6 | 15.5 | 5.5 | -52.5 | 3.45 | 1.11 | -2.68 | 5.14 |
| $\kappa = 1.0$ | 28 | 45.8 | 6.3 | 2.0 | -84.8 | 6.22 | 6.21 | -5.48 | 5.97 |
| Benchmark | 26 | 49.7 | 10.8 | 3.6 | -62.9 | 4.65 | 3.09 | -3.94 | 5.66 |

the economy. These parameters are a discrete approximation for a continuous process for labor income, $li_t \equiv we_t n_t$, that is

$$\log\left(li_{t+1}\right) = \rho \log\left(li_t\right) + \varepsilon, \quad \text{where } \varepsilon \sim N\left(0, \sigma_\varepsilon^2\right).$$

In our benchmark calibration we target $\rho = 0.9$ and $\sigma_\varepsilon = 0.2$. Given the importance of these choices for our results and the lack of consensus in the literature about them (see Section 3.4 for a discussion), we provide here the results for alternative numbers for $\rho$ and $\sigma_\varepsilon$. For each of these we recalibrate the economy modifying only the corresponding target, Table 10 contains the results.

As one would expect, the magnitudes of the results do change considerably given changes in these important parameters. However, reassuringly, the qualitative features of the fiscal policy instruments and of where the welfare gains come from is not substantially affected.

## 7.2 Labor supply and intertemporal elasticities

One parameter, $\sigma$, determines three important aspects of our benchmark experiment: the agents' intertemporal elasticity of substitution and relative risk aversion, and the planner's degree of inequality aversion. Table 10 contains the results for other choices of this parameter and also for different levels of Frisch elasticity.

When $\sigma$ is reduced from 2 to 1, the planner's inequality aversion is reduced and, accordingly, capital income taxes are kept at the upper bound for less periods ($t^*$ goes from 26 to 12). Moreover, the agents' intertemporal elasticity of substitution increases and their risk aversion is reduced which implies that long-run capital taxes lead to, at the same time, higher distortions and less benefits. It follows that the optimal long-run capital tax is lower. This leads to a higher proportion of welfare gains coming from the level effect and less coming from redistribution. The opposite happens when $\sigma$ is increased to $3$. Intuitively, a higher Frisch elasticity implies a lower optimal labor income tax and a higher associated level effect. Notice that these results are in line with the propositions established in Section 1.

# 8 Conclusion

In this paper we quantitatively characterize the solution to the Ramsey problem in the standard incomplete markets model. We find that even though the planner has the ability to obtain all revenue via non distortive lump-sum taxes, it chooses instead to tax capital income heavily and labor income to a lesser extent. Moreover, we show that it is beneficial for the government to accumulate assets over time. By decomposing the welfare gains we diagnose that, relative to the current US tax system, this policy leads to an overall reduction of the distortions of agent's decisions, to a substantial amount of redistribution and to a reduction in the amount of insurance provided by the government. Importantly, we also show that disregarding the transitory dynamics and focusing only on steady states can lead to severely misleading results.

Finally, we do not view our results as a final answer to our initial question: to what extent should governments use fiscal policy instruments to provide redistribution and insurance? Instead, we understand it as a contribution to the debate. The model we use abstracts from important aspects of reality, as any useful model must, and we miss some important dimensions. For instance, in the model studied above an agent's productivity is entirely a matter of luck, it would be interesting to understand the effects of allowing for human capital accumulation. We also assume the government has the ability to fully commit to future policies, relaxing this assumption could lead to interesting insights.

# References

**Acikgoz, Omer**, "Transitional Dynamics and Long-run Optimal Taxation Under Incomplete Markets," MPRA Paper 50160, University Library of Munich, Germany December 2015.

**Aiyagari, S Rao**, "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, August 1994, *109* (3), 659–684.

__ , "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, December 1995, *103* (6), 1158–1175.

**Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala**, "Optimal Taxation without State-Contingent Debt," *Journal of Political Economy*, December 2002, *110* (6), 1220–1254.

— **and Ellen R. McGrattan**, "The optimum quantity of debt," *Journal of Monetary Economics*, October 1998, *42* (3), 447–469.

**Atkeson, Andrew, V.V. Chari, and Patrick J. Kehoe**, "Taxing capital income: a bad idea," *Quarterly Review*, 1999, (Sum), 3–17.

**Bassetto, Marco**, "Optimal fiscal policy with heterogeneous agents," *Quantitative Economics*, November 2014, *5* (3), 675–704.

**Benabou, Roland**, "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, March 2002, *70* (2), 481–517.

**Bewley, Truman**, "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," in Werner Hildenbrand and Andreu Mas-Collel, eds., *Contributions to Mathematical Economics in Honor of Gerard Debreu*, Amsterdam: North-Holland, 1986.

**Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas Sargent**, "Fiscal Policy and Debt Management with Incomplete Markets," *Quarterly Journal of Economics*, 2016.

**Castañeda, Ana, Javier Díaz-Giménez, and José-Víctor Ríos-Rull**, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, August 2003, *111* (4), 818–857.

**Chamley, Christophe**, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, May 1986, *54* (3), 607–22.

— , "Capital income taxation, wealth distribution and borrowing constraints," *Journal of Public Economics*, January 2001, *79* (1), 55–69.

**Chatterjee, Satyajit**, "Transitional dynamics and the distribution of wealth in a neoclassical growth model," *Journal of Public Economics*, May 1994, *54* (1), 97–119.

**Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber**, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, May 2011, *101* (3), 471–75.

— **,** — **,** — **, and** — , "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities," in "NBER Macroeconomics Annual 2012, Volume 27" NBER Chapters, National Bureau of Economic Research, Inc, 2012, pp. 1–56.

**Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger**, "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, March 2009, *99* (1), 25–48.

**Dávila, Julio, Jay H. Hong, Per Krusell, and José-Víctor Ríos-Rull**, "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks," *Econometrica*, November 2012, *80* (6), 2431–2467.

**Díaz-Giménez, Javier, Andrew Glover, and José-Víctor Ríos-Rull**, "Facts on the distributions of earnings, income, and wealth in the United States: 2007 update," *Quarterly Review*, 2011.

**Domeij, David and Jonathan Heathcote**, "On The Distributional Effects Of Reducing Capital Taxes," *International Economic Review*, 05 2004, *45* (2), 523–554.

**Erosa, Andres and Martin Gervais**, "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, August 2002, *105* (2), 338–369.

**Evans, David**, "Optimal Taxation with Persistent Idiosyncratic Investment Risk," mimeo November 2014.

**Farhi, Emmanuel and Iván Werning**, "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation," *Journal of Political Economy*, 2012, *120* (3), 000 – 000.

**_ and _** , "Insurance and Taxation over the Life Cycle," *Review of Economic Studies*, 2013, *80* (2), 596–635.

**Floden, Martin**, "The effectiveness of government debt and transfers as insurance," *Journal of Monetary Economics*, August 2001, *48* (1), 81–108.

**Golosov, Michael, Narayana Kocherlakota, and Aleh Tsyvinski**, "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 2003, *70* (3), 569–587.

**Gottardi, Piero, Atsushi Kajii, and Tomoyuki Nakajima**, "Constrained Inefficiency and Optimal Taxation with Uninsurable Risks," KIER Working Papers 694, Kyoto University, Institute of Economic Research August 2014.

**_ , _ , and _** , "Optimal Taxation and Debt with Uninsurable Risks to Human Capital Accumulation," *American Economic Review*, November 2015, *105* (11), 3443–70.

**Greenwood, Jeremy, Zvi Hercowitz, and Gregory W Huffman**, "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, June 1988, *78* (3), 402–17.

**Guvenen, Fatih and Anthony Smith**, "Inferring labor income risk and partial insurance from economic choices," Staff Report 485, Federal Reserve Bank of Minneapolis 2013.

**Hagedorn, Marcus, Hans Holter, and Yikai Wang**, "Optimum Quantity of Capital and Debt," Technical Report February 2016.

**Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante**, "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 08 2010, *118* (4), 681–722.

**_ , _ , and _** , "Optimal Tax Progressivity: An Analytical Framework," NBER Working Papers 19899, National Bureau of Economic Research, Inc February 2014.

**Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen**, "The Carnegie Conjecture: Some Empirical Evidence," *The Quarterly Journal of Economics*, 1993, *108* (2), pp. 413–435.

**Huggett, Mark**, "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 953–969.

**Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote**, "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review*, September 2001, *91* (4), 778–794.

**Imrohoruglu, Ayse**, "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, December 1989, *97* (6), 1364–83.

**Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi**, "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, March 1997, *73* (1), 93–117.

**Judd, Kenneth L.**, "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics*, October 1985, *28* (1), 59–83.

**Krueger, Dirk and Alexander Ludwig**, “Optimal Progressive Taxation and Education Subsidies in a Model of Endogenous Human Capital Formation,” PIER Working Paper Archive 13-035, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania March 2013.

— **and Fabrizio Perri**, “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *Review of Economic Studies*, 2006, *73* (1), 163–193.

**Laczo, Sarolta, Albert Marcet, and Katharina Greulich**, “Pareto-Improving Optimal Capital and Labor Taxes,” Technical Report 2015.

**Marcet, Albert, Francesc Obiols-Homs, and Philippe Weil**, “Incomplete markets, labor supply and capital accumulation,” *Journal of Monetary Economics*, November 2007, *54* (8), 2621–2635.

**Nakajima, Makoto**, “Optimal capital income taxation with housing,” Working Papers 10-11, Federal Reserve Bank of Philadelphia 2010.

— , “Business Cycles In The Equilibrium Model Of Labor Market Search And Self-Insurance,” *International Economic Review*, 05 2012, *53* (2), 399–432.

**Panousi, Vasia and Catarina Reis**, “Optimal capital taxation with idiosyncratic investment risk,” Technical Report 2012.

**Poschke, Markus, Baris Kaymak, and Ozan Bakis**, “On the Optimality of Progressive Income Redistribution,” 2012 Meeting Papers 837, Society for Economic Dynamics 2012.

**Straub, Ludwig and Iván Werning**, “Positive Long Run Capital Taxation: Chamley-Judd Revisited,” NBER Working Papers 20441, National Bureau of Economic Research, Inc August 2014.

**Trabandt, Mathias and Harald Uhlig**, “The Laffer curve revisited,” *Journal of Monetary Economics*, 2011, *58* (4), 305–327.

**Troshkin, Maxim, Aleh Tsyvinski, and Mikhail Golosov**, “Optimal Dynamic Taxes,” 2010 Meeting Papers 320, Society for Economic Dynamics 2010.

**Werning, Ivn**, “Optimal Fiscal Policy with Redistribution,” *The Quarterly Journal of Economics*, 08 2007, *122* (3), 925–967.

**Winter, Christoph and Sigrid Roehrs**, “Reducing Government Debt in the Presence of Inequality,” 2014 Meeting Papers 176, Society for Economic Dynamics February 2016.

**Wolff, Edward Nathan**, “Recent Trends in Household Wealth, 1983-2009: The Irresistible Rise of Household Debt,” *Review of Economics and Institutions*, 2011, *2* (1).

# Appendix

This appendix presents concise versions of the proofs. Extensive versions with more details are contained in a separate Online Appendix which can be found in our websites.[41]

## A Proofs for two-period economies

### A.1 Uncertainty economy

Define $\tau_R^k \equiv r\tau^k/(1+r)$. Six equations determine a tax distorted equilibrium $(K, n_L, n_H, r, w;$ $\tau^n$, $\tau_R^k, T)$ according to Definition 1: the first order conditions of the agent's problem (one intertemporal and two intratemporal), the first order conditions of the firm's problem

$$r = f_K(K,N), \quad \text{and} \quad w = f_N(K,N), \quad \text{where } N = \pi e_L n_L + (1-\pi) e_H n_H \tag{8.1}$$

and the government's budget constraint. Using equation (8.1) to substitute out for $r$ and $w$ we are left with a system of four equations that any vector $\left(K, n_L, n_H, \tau^n, \tau_R^k, T\right)$ of equilibrium values must satisfy. The two degrees of freedom are a result of the fact that the planner has three instruments $\left(\tau^n, \tau_R^k, T\right)$ that are restricted by one equation, the government's budget constraint. Defining welfare by

$$W \equiv u(\omega - K, \bar{n}) + \beta E\left[u\left((1-\tau^n) f_N(K,N) e_i n_i + \left(1-\tau_R^k\right) f_K(K,N) K + T\right), n_i\right]$$

and totally differentiating the four equilibrium equations together with this definition and making the appropriate simplifications using Assumption 1 we obtain the following equation (the algebra is tedious and, therefore, suppressed[42]):

$$dW = \Theta^n d\tau^n + \Theta^k d\tau_R^k,$$

where $\Theta^n$ and $\Theta^k$ are complicated functions of equilibrium variables[43].

**Lemma 2** *Under Assumption 1, in equilibrium $n_H > n_L$ and $u_c(c_L, n_L) > u_c(c_H, n_H)$.*

The proof of this Lemma is contained in the Online Appendix.

**Proof of Proposition 1.** First notice that the optimal tax system must satisfy $\Theta^n = 0$ and $\Theta^k = 0$, otherwise there would exist variations in $\left(\tau^n, \tau_R^k\right) \in (-\infty, 1)^2$ that would increase welfare. $\Theta^k =$

[41] http://www.dyrda.info/ or http://sites.google.com/site/marcelozouainpedroni/

[42] Mathematica codes that compute all the algebraic steps are available in our websites.

[43] Here are the exact formulas:

$$\Theta^k \equiv \frac{f_K K U_c}{\Phi}\left\{f_N f_{KN} N\left[(1-\tau^n)(V_c - U_c) + \tau^n \kappa U_c\right] + \tau_R^k f_K (f_N + f_{KN} K\kappa) U_c\right\}.$$

$$\begin{aligned}\Theta^n \equiv \frac{f_N N}{(1-\tau^n)\Phi}&\left\{\left(1-\tau_R^k\right) f_K^2 f_N K\left[(1-\tau^n)\left(U_{cc}(U_c - V_c) + \tau_R^k (V_{cc} - U_{cc}) U_c\right) - \left(1-\tau_R^k\right)\tau^n \kappa U_{cc} U_c\right]\right.\\ &\left.+ f_N\left[(1-\tau^n)(V_c - U_c) + \tau^n \kappa U_c\right]\left[\left(1-\tau_R^k\right) f_{KN} N U_c - K u_{cc}^0\right] + \left(1-\tau_R^k\right)\tau_R^k f_{KN} f_K K \kappa U_c^2\right\}.\end{aligned}$$

0 simplifies to $\theta_1^k + \theta_2^k \tau^n + \theta_3^k \tau_R^k = 0$ where

$$\theta_1^k \equiv f_N f_{KN} N \left(V_c - U_c\right), \theta_2^k \equiv f_N f_{KN} N \left(\left(1+\kappa\right) U_c - V_c\right), \text{ and } \theta_3^k \equiv f_K \left(f_N + \kappa f_{KN} K\right) U_c.$$

Solving this equation for $\tau_R^k$, substituting it in $\Theta^n = 0$ and simplifying entails

$$V_c \left(1 - \tau^n\right) - U_c \left(1 - \left(1+\kappa\right) \tau^n\right) = 0.$$

Solving for $\tau^n$ we obtain equation (1.1) and substituting it back in the equation for $\tau_R^k$ we obtain $\tau_R^k = 0$; and, therefore, $\tau^k = 0$. This is the only pair $\left(\tau^n, \tau_R^k\right) \in \left(-\infty, 1\right)^2$ that solves the system $\Theta^n = 0$ and $\Theta^k = 0$. The fact that the optimal level of $\tau^n > 0$ follows from Lemma 2. ■

### A.2 Inequality economy

The proof of Proposition 2 is entirely analogous and for that reason suppressed here. It can be found in the Online Appendix.

## B Welfare decomposition

Let $v\left(x\right) \equiv u\left(c, n\right)$ where $x$ is the consumption-labor composite defined in Section 4.4 and $u$ is defined in (3.1). Consider a policy reform. Denote by $x_t^R\left(a_0, e^t\right)$ the equilibrium consumption-labor composite path of an agent with initial assets $a_0$ and history of productivities $e^t$ if the reform is implemented. Let $x_t^{NR}\left(a_0, e^t\right)$ be the equilibrium path in case there is no reform. The average welfare gain, $\Delta$, that results from implementing the reform is defined as the constant percentage increase to $x_t^{NR}\left(a_0, e^t\right)$ that equalizes the (utilitarian) welfare to the value associated with the reform, that is,

$$\int E_0\left[U\left(\left(1+\Delta\right)\left\{x_t^{NR}\left(a_0, e^t\right)\right\}\right)\right] d\lambda_0\left(a_0, e_0\right) = \int E_0\left[U\left(\left\{x_t^{R}\left(a_0, e^t\right)\right\}\right)\right] d\lambda_0\left(a_0, e_0\right), \tag{8.2}$$

where $\lambda_0$ is the initial distribution over states $\left(a_0, e_0\right)$ and $U\left(\left\{x_t\left(a_0, e^t\right)\right\}\right) \equiv \sum_{t=0}^{\infty} \beta^t v(x_t(a_0, e^t)) = \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(a_0, e^t\right), n_t\left(a_0, e^t\right)\right)$.

---

where

$$U_c \equiv \beta\left[\pi u_c\left(c_L, n_L\right) + \left(1-\pi\right) u_c\left(c_H, n_H\right)\right], \quad U_{cc} \equiv \beta\left[\pi u_{cc}\left(c_L, n_L\right) + \left(1-\pi\right) u_{cc}\left(c_H, n_H\right)\right],$$
$$V_c \equiv \beta\left[\pi u_c\left(c_L, n_L\right) \frac{e_L n_L}{N} + \left(1-\pi\right) u_c\left(c_H, n_H\right) \frac{e_H n_H}{N}\right],$$
$$V_{cc} \equiv \beta\left[\pi u_{cc}\left(c_L, n_L\right) \frac{e_L n_L}{N} + \left(1-\pi\right) u_{cc}\left(c_H, n_H\right) \frac{e_H n_H}{N}\right],$$
$$\Phi \equiv \left(1-\tau_R^k\right)\left(f_K f_N f_{KN} K N\left(\left(1-\tau^n\right)\left(V_{cc} - U_{cc}\right) + \tau^n \kappa U_{cc}\right) + \left(f_N + f_{KN} K \kappa\right) f_K^2 K U_{cc} - f_N f_{KN} N U_c\right) + \left(f_N + f_{KN} K \kappa\right) K u_{cc}^0.$$

Define

$$X_t^j \equiv \int x_t^j \left(a_0, e^t\right) d\lambda_t^j \left(a_0, e^t\right), \quad \text{for } j = R, NR.$$

to be the average level of $x$ at each $t$. Then, the level effect, $\Delta_L$, is

$$U\left(\left(1+\Delta_L\right)\left\{X_t^{NR}\right\}\right) = U\left(\left\{X_t^R\right\}\right), \tag{8.3}$$

In order to define the other two components we need some previous definitions. Let $\bar{x}^j\left(a_0, e_0\right)$ denote the individual consumption-labor certainty equivalent,

$$U\left(\left\{\bar{x}^j\left(a_0, e_0\right)\right\}\right) = E_0\left[U\left(\left\{x_t^j\left(a_0, e^t\right)\right\}\right)\right], \quad \text{for } j = R, NR, \tag{8.4}$$

(notice that $\bar{x}^j\left(a_0, e_0\right)$ can be chosen to be constant) and let $\bar{X}^j$ be the aggregate consumption-labor certainty equivalent,

$$\bar{X}^j = \int \bar{x}^j\left(a_0, e_0\right) d\lambda\left(a_0, e_0\right), \quad \text{for } j = R, NR. \tag{8.5}$$

The insurance effect, $\Delta_I$, is defined by

$$1+\Delta_I \equiv \frac{1-p_{unc}^R}{1-p_{unc}^{NR}}, \text{ where } \quad U\left(\left(1-p_{unc}^j\right)\left\{X_t^j\right\}\right) = U\left(\left\{\bar{X}^j\right\}\right), \tag{8.6}$$

and the redistribution effect, $\Delta_R$, by

$$1+\Delta_R \equiv \frac{1-p_{ine}^R}{1-p_{ine}^{NR}}, \text{ where } \quad U\left(\left(1-p_{ine}^j\right)\left\{\bar{X}^j\right\}\right) = \int U\left(\left\{\bar{x}^j\left(a_0, e_0\right)\right\}\right) d\lambda\left(a_0, e_0\right). \tag{8.7}$$

The following proposition holds[44].

**Proof of Proposition 3.** First notice that $v(x) \equiv u(c, n)$ where $u$ is the GHH utility function, defined in (3.1), satisfies the following regularity property: there exists a totally multiplicative function $h$ : (i.e. $h(ab) = h(a)h(b)$, and $h(a/b) = h(a)/h(b)$) such that for any scalar $\alpha$,

$$v(\alpha x) = h(\alpha) v(x). \tag{8.8}$$

[44]This result is similar to the one introduced by Benabou (2002) and used in Floden (2001).

Hence, suppressing the dependence on $(a_0, e_0)$, we obtain:

$$
\begin{aligned}
\int E_0 U\left(\left\{x_t^R\right\}\right) d\lambda_0^R &\overset{(8.4)}{=} \int U\left(\left\{\bar{x}^R\right\}\right) d\lambda_0^R \overset{(8.7)}{=} U\left(\left(1-p_{ine}^R\right)\left\{\bar{X}^R\right\}\right) \overset{(8.8)}{=} h\left(1-p_{ine}^R\right) U\left(\left\{\bar{X}^R\right\}\right) \\
&\overset{(8.6)}{=} h\left(1-p_{ine}^R\right) U\left(\left(1-p_{unc}^R\right)\left\{X_t^R\right\}\right) \overset{(8.8)}{=} h\left(\left(1-p_{ine}^R\right)\left(1-p_{unc}^R\right)\right) U\left(\left\{X_t^R\right\}\right) \\
&\overset{(8.3)}{=} h\left(\left(1-p_{ine}^R\right)\left(1-p_{unc}^R\right)\right) U\left(\left(1+\Delta_L\right)\left\{X_t^{NR}\right\}\right) \\
&\overset{(8.8)}{=} h\left(\left(1+\Delta_L\right)\left(1-p_{ine}^R\right)\left(1-p_{unc}^R\right)\right) U\left(\left\{X_t^{NR}\right\}\right) \\
&\overset{(8.8)}{=} h\left(\left(1+\Delta_L\right)\left(1-p_{ine}^R\right)\frac{\left(1-p_{unc}^R\right)}{\left(1-p_{unc}^{NR}\right)}\right) U\left(\left(1-p_{unc}^{NR}\right)\left\{X_t^{NR}\right\}\right) \\
&\overset{(8.6)}{=} h\left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\left(1-p_{ine}^R\right)\right) U\left(\left\{\bar{X}^{NR}\right\}\right) \\
&\overset{(8.8)}{=} h\left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\frac{\left(1-p_{ine}^R\right)}{\left(1-p_{ine}^{NR}\right)}\right) U\left(\left(1-p_{ine}^{NR}\right)\left\{\bar{X}^{NR}\right\}\right) \\
&\overset{(8.7)}{=} h\left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\left(1+\Delta_R\right)\right) \int U\left(\left\{\bar{x}^{NR}\right\}\right) d\lambda_0^{NR} \\
&\overset{(8.6)}{=} h\left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\left(1+\Delta_R\right)\right) \int E_0 U\left(\left\{x_t^{NR}\right\}\right) d\lambda_0^{NR} \\
&\overset{(8.8)}{=} \int E_0 U\left(\left(1+\Delta_R\right)\left(1+\Delta_I\right)\left(1+\Delta_L\right)\left\{x_t^{NR}\right\}\right) d\lambda_0^{NR}.
\end{aligned}
$$

The result follows from the definition of $\Delta$ in equation (8.2). ■

## C Proofs for complete market economies

The proofs follow straight-forwardly the approach introduced by Werning (2007). Hence, for details on the logic behind the procedure we refer the reader to Online Appendix , here we focus mainly on the parts that comprise our value added. We depart from Werning (2007) in following ways: we use the GHH utility function (whereas he studies the separable and Cobb-Douglas cases), we do not allow the Ramsey planner to choose time zero policies and impose an upper bound of $1$ for capital income taxes. These departures make the Ramsey planner's problem comparable to our benchmark experiment. The restriction on time zero policies is particularly important because it prevents the planner from confiscating the (potentially unequal) initial capital levels eliminating the corresponding redistribution motives.

Consider Economy 4 as described in Section 5. For simplicity, we assume that agents are divided into a finite number of types $i \in I$ of relative size $\pi_i$. Type $i$ has an initial asset position of $a_{i,0}$ and a productivity level of $e_i$. Let $p_t$ denote the price of the consumption good in period $t$ in terms of period $0$. Since markets are complete we can write down the present value budget constraint of the agent (remember that $\tau^c$ is a parameter),

$$\sum_{t=0}^{\infty} p_t \left( (1+\tau^c) c_{i,t} + a_{i,t+1} \right) \leq \sum_{t=0}^{\infty} p_t \left( (1-\tau_t^n) w_t e_i n_{i,t} + R_t a_{i,t} + T_t \right),$$

where $R_t \equiv 1 + \left(1 - \tau_t^k\right) r_t$. Rule out arbitrage opportunities by setting $p_t = R_{t+1} p_{t+1}$, and define $T \equiv \sum_{t=0}^{\infty} p_t T_t$. Then, the budget constraint simplifies to

$$\sum_{t=0}^{\infty} p_t \left( (1+\tau^c) c_{i,t} - (1-\tau_t^n) w_t e_i n_{i,t} \right) \leq R_0 a_{i,0} + T. \tag{8.9}$$

Similarly, the government's budget constraint simplifies to

$$R_0 B_0 + T + \sum_t p_t G = \sum_t p_t \left( \tau^c C_t + \tau_t^n w_t N_t + \tau_t^k r_t K_t \right). \tag{8.10}$$

The resource constraint is given by

$$C_t + G + K_{t+1} = f(K_t, N_t), \quad \text{for all } t \geq 0. \tag{8.11}$$

**Definition 6** *Given $\{a_{i,0}\}$, $K_0$, $B_0$ and $\left(\tau_0^n, \tau_0^k, T_0\right)$, a competitive equilibrium is a policy $\{\tau_t^n, \tau_t^k, T_t\}_{t=1}^{\infty}$, a price system $\{p_t, w_t, r_t\}_{t=0}^{\infty}$, and an allocation $\{c_{i,t}, n_{i,t}, K_{t+1}\}_{t=0}^{\infty}$, such that: (i) agents choose $\{c_{i,t}, n_{i,t}\}_{t=0}^{\infty}$ to maximize utility subject to budget constraint (8.9) taking policies and prices (that satisfy $p_t = R_{t+1} p_{t+1}$) as given; (ii) firms maximize profits; (iii) the government's budget constraint (8.10) holds; and (iv) markets clear: the resource constraints (8.11) hold.*

Given aggregate levels $C_t$ and $N_t$, individual consumption and labor supply levels can be found by solving the following static subproblem

$$U(C_t, N_t; \varphi) \equiv \max_{c_{i,t}, n_{i,t}} \sum_i \pi_i \varphi_i u(c_{i,t}, n_{i,t}) \quad \text{s.t.} \sum_i \pi_i c_{i,t} = C_t \quad \text{and} \quad \sum_i \pi_i e_i n_{i,t} = N_t \tag{8.12}$$

where $u$ is given by equation (3.1), for some vector $\varphi \equiv \{\varphi_i\}$ of market weights $\varphi_i \geq 0$. Let $c_{i,t}^m(C_t, N_t; \varphi)$, and $n_{i,t}^m(C_t, N_t; \varphi)$ be the $\arg\max$ of this problem. It can be shown that[45]

[45]Where constants are defined as follows:

$$\omega_i^c \equiv \frac{(\varphi_i)^{\frac{1}{\sigma}}}{\sum_j \pi_j (\varphi_j)^{\frac{1}{\sigma}}}, \quad \omega_i^n \equiv \frac{(e_i)^{\kappa}}{\sum_j \pi_j (e_j)^{1+\kappa}}, \quad \Omega^c \equiv \left( \sum_i \pi_i (\varphi_i)^{\frac{1}{\sigma}} \right)^{\sigma}, \quad \text{and} \quad \Omega^n \equiv \left( \sum_j \pi_j (e_j)^{1+\kappa} \right)^{-\frac{1}{\kappa}}$$

$$
\begin{aligned}
c_{i,t}^{m}\left(C_t, N_t; \varphi\right) &= \omega_i^c C_t + \chi \frac{\kappa}{1+\kappa}\left(\left(\omega_i^n\right)^{\frac{1+\kappa}{\kappa}} - \omega_i^c \Omega^n\right)\left(N_t\right)^{\frac{1+\kappa}{\kappa}} \\
n_{i,t}^{m}\left(C_t, N_t; \varphi\right) &= \omega_i^n N_t \\
U\left(C_t, N_t; \varphi\right) &= \frac{\Omega^c}{1-\sigma}\left(C_t - \Omega^n \chi \frac{\kappa}{1+\kappa}\left(N_t\right)^{\frac{1+\kappa}{\kappa}}\right)^{1-\sigma}
\end{aligned}
$$

Then, implementability constraints can be written as

$$
\begin{aligned}
&\sum_{t=0}^{\infty} \beta^t (U_C\left(C_t, N_t; \varphi\right) c_{i,t}^{m}\left(C_t, N_t; \varphi\right) + U_N\left(C_t, N_t; \varphi\right) e_i n_{i,t}^{m}\left(C_t, N_t; \varphi\right)) \\
&\qquad\qquad = U_C\left(C_0, N_0; \varphi\right)\left(\frac{R_0 a_{i,0} + T}{1+\tau^c}\right) \quad \text{for all } i \in I
\end{aligned} \tag{8.13}
$$

**Proposition 7** *An aggregate allocation $\{C_t, N_t, K_{t+1}\}_{t=0}^{\infty}$ can be supported by a competitive equilibrium if and only if the resource constraints (8.11) hold and there exist market weights $\varphi$ and a lump-sum tax $T$ so that the implementability conditions (8.13) hold for all $i \in I$. Individual allocations can then be computed using functions $c_{i,t}^{m}$ and $n_{i,t}^{m}$, prices and taxes can be computed using the usual equilibrium conditions.*

The Ramsey problem is that of choosing policies $\left\{\tau_t^n, \tau_t^k, T_t\right\}_{t=1}^{\infty}$, taking $\{a_{i,0}\}$, $K_0$, $B_0$ and $\left(\tau_0^n, \tau_0^k, T_0\right)$ as given, to maximize a weighted sum of the individual utilities,

$$
\sum_{t=0}^{\infty} \beta^t \pi_i \lambda_i u\left(c_{i,t}, n_{i,t}\right), \tag{8.14}
$$

where $\{\lambda_i\}$ are the welfare weights normalized so that $\sum_i \pi_i \lambda_i = 1$ with $\lambda_i \geq 0$, subject to allocations and policies being a part of a competitive equilibrium and $\tau_t^k \leq 1$ for all $t \geq 1$.

First notice that in equilibrium it must be that $U_C(t) = \beta\left(1 + \left(1 - \tau_{t+1}^k\right) r_{t+1}\right) U_C(t+1)$, so that

$$
U_C(t) \geq \beta U_C(t+1), \tag{8.15}
$$

is equivalent to $\tau_{t+1}^k \leq 1$. Moreover, notice that $\tau_0^k$ and $T_0$ have not been substituted out in the implementability constraint. The fact that $\tau_0^n$ is given together with the equilibrium condition $\left(1-\tau_0^n\right) w_0 = -U_N(0)/U_C(0)$ is equivalent to

$$
N_0 = \bar{N}_0, \tag{8.16}
$$

where $\bar{N}_0$ is defined implicitly as a function of variables given to the Ramsey planner,

$$
\left(1-\tau_0^n\right) f_N\left(K_0, \bar{N}_0\right) = \Omega^n \chi \left(\bar{N}_0\right)^{\frac{1}{\kappa}}.
$$

Finally, we can use Proposition 7 to rewrite the Ramsey problem as that of choosing $\{C_t, N_{t+1}, K_{t+1}\}_{t=0}^{\infty}$, $T$, and $\varphi$ to maximize (8.14) subject to (8.11) for all $t \geq 0$, (8.13) for all $i \in I$ with multiplier $\mu_i$, (8.15) for all $t \geq 0$ with multiplier $\eta_t$, and (8.16). Equivalently, we can write it as that of solving the following auxiliary problem

$$\max_{\{C_t,N_{t+1},K_{t+1}\}_{t=0}^{\infty},T,\varphi} \sum_{t=0}^{\infty} \beta^t W\left(C_t, N_t; \varphi, \mu, \lambda\right) - U_C\left(C_0, N_0; \varphi\right) \sum_{i\in I} \pi_i \mu_i \left(\frac{R_0 a_{i,0} + T}{1+\tau^c}\right),$$

subject to (8.11) for all $t \geq 0$, (8.15) for all $t \geq 0$, and (8.16), where

$$\begin{aligned} W\left(C_t, N_t; \varphi, \mu, \lambda\right) \equiv \sum_i \pi_i \{ & \lambda_i u\left(c_{i,t}^m\left(C_t, N_t; \varphi\right), n_{i,t}^m\left(C_t, N_t; \varphi\right)\right) \\ & + \mu_i \left(U_C\left(C_t, N_t; \varphi\right) c_{i,t}^m\left(C_t, N_t; \varphi\right) + U_N\left(C_t, N_t; \varphi\right) e_i n_{i,t}^m\left(C_t, N_t; \varphi\right)\right)\}. \end{aligned}$$

With some algebra it can be shown that[46]

$$\begin{aligned} W\left(C_t, N_t; \varphi, \mu, \lambda\right) = & \frac{1}{1-\sigma}\left(C_t - \Omega^n \chi \frac{\kappa}{1+\kappa}\left(N_t\right)^{\frac{1+\kappa}{\kappa}}\right)^{-\sigma} \\ & * \left(\Phi C_t - \left(\Phi + (1-\sigma)\Psi\right)\Omega^n \chi \frac{\kappa}{1+\kappa}\left(N_t\right)^{\frac{1+\kappa}{\kappa}}\right) \end{aligned} \tag{8.17}$$

Define $R_t^* \equiv 1 + r_t$ and

$$\eta_{-1} \equiv \frac{R_0}{\beta\left(1+\tau^c\right)} \sum_i \pi_i \mu_i a_{i,0},$$

and first order conditions (for the following proofs we need only necessary conditions) together with equilibrium conditions imply the following equations[47]

$$\sum_i \pi_i \mu_i = 0 \tag{8.18}$$

$$\frac{\tau_t^n + \tau^c}{1+\tau^c} = \frac{\Psi \Theta_t}{\Phi \Theta_t + \Psi\left(\sigma + \Theta_t\right) + \Upsilon_t \sigma\left(\beta \eta_{t-1} - \eta_t\right)}, \text{ for } t \geq 1 \tag{8.19}$$

$$\frac{R_{t+1}}{R_{t+1}^*} = \frac{\Phi \Theta_{t+1} + \Psi \sigma + \Upsilon_{t+1} \sigma\left(\beta \eta_t - \eta_{t+1}\right)}{\Phi \Theta_t + \Psi \sigma + \Upsilon_t \sigma\left(\beta \eta_{t-1} - \eta_t\right)} \frac{\Theta_t}{\Theta_{t+1}}, \text{ for } t \geq 0 \tag{8.20}$$

[46] Where constants are defined as follows:

$$\Phi \equiv \left(\Omega^c\right)^{\frac{\sigma-1}{\sigma}} \sum_i \pi_i (\varphi_i)^{\frac{1}{\sigma}} \left(\frac{\lambda_i}{\varphi_i} + (1-\sigma)\mu_i\right), \text{ and } \Psi \equiv \frac{\Omega^c}{\kappa} \sum_j \pi_j \mu_j e_j \omega_j^n.$$

[47] Where $\Upsilon_t \equiv \Omega^c/(1-\sigma)\Omega^n \chi \frac{\kappa}{1+\kappa}\left(N_t\right)^{\frac{1+\kappa}{\kappa}}$.

Notice that $\Upsilon_t > 0$ and $\Theta_t > 0$, for all $t \geq 0$.

### C.1 Economy 2

**Lemma 3** *If $e_i = 1$ for all $i \in I$, then $\Psi = 0$ and $\Phi > 0$.*

**Proof.** If $e_i = 1$ for all $i \in I$, then it follows from the definition of $\Psi$ that

$$\Psi = \frac{\Omega^c}{\kappa} \frac{\sum_j \pi_j \mu_j (e_j)^{1+\kappa}}{\sum_j \pi_j (e_j)^{1+\kappa}} = \frac{\Omega^c}{\kappa} \frac{\sum_j \pi_j \mu_j}{\sum_j \pi_j} = 0,$$

where the last equality follows from equation (8.18). Since $\Psi = 0$, it follows from equation (8.17) that

$$W(C_t, N_t; \varphi, \mu, \lambda) = \frac{\Phi}{1-\sigma} \left( C_t - \Omega^n \chi \frac{\kappa}{1+\kappa} (N_t)^{\frac{1+\kappa}{\kappa}} \right)^{1-\sigma}.$$

If $\Phi \leq 0$ it would be optimal to set $C_t = 0$ for all $t \geq 0$ which cannot be a solution to the initial Ramsey problem. ■

**Proof of Proposition 4.** Using Lemma 3, from equation (8.19) it follows that

$$\tau_t^n = -\tau^c, \text{ for } t \geq 1.$$

Next, suppose $\eta_t = 0$, for all $t \geq 0$. Then, it follows from (8.20) that $\tau_1^k < 1$ if

$$-\frac{1}{\beta} \frac{\Phi\Theta_0}{\Upsilon_0 \sigma} \equiv P_1 < \eta_{-1} < M_1 \equiv \frac{1}{\beta} \frac{(R_1^* - 1)\Phi\Theta_0}{\Upsilon_0 \sigma},$$

and that $\tau_t^k = 0$ for $t \geq 2$. Hence, if $P_1 < \eta_{-1} < M_1$, the constraints will in fact never be binding. Now, suppose $\eta_t > 0$, for $t \leq t^* - 2$ and $\eta_t = 0$, for all $t \geq t^* - 1$, then it follows from (8.20) that $\tau_{t^*}^k < 1$ if

$$-\sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\Phi\Theta_{\tau-1}}{\Upsilon_{\tau-1}\sigma} \equiv P_{t^*} < \eta_{-1} < M_{t^*} \equiv \sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\left(\prod_{t=\tau}^{t^*} R_t^* - 1\right)\Phi\Theta_{\tau-1}}{\Upsilon_{\tau-1}\sigma},$$

and that $\tau_t^k = 0$ for $t \geq t^*+1$. The result follows from the fact that $\eta_{-1}$ is finite, $\lim_{t\to\infty} P_t = -\infty$ and $\lim_{t\to\infty} M_t = \infty$. ■

## C.2 Economy 3

**Proof of Proposition 5.** In this economy there is no heterogeneity in initial levels of asset, i.e. $a_{i,0} = a_0$ for all $i \in I$. Then it follows that

$$\eta_{-1} = \frac{R_0}{\beta(1+\tau^c)} \sum_i \pi_i \mu_i a_{i,0} = \frac{R_0}{\beta(1+\tau^c)} a_0 \sum_i \pi_i \mu_i = 0$$

where the last equality follows from equation (8.18). Since here we assume that $\tau_t^k$ does not have to be bounded by 1, it follows that $\eta_t = 0$ for all $t \geq 1$. Then, equation (5.1) follows directly from equation (8.19), (5.2) from its derivative with respect to $\Theta_t$, and (5.3) from equations (8.19) and (8.20). ■

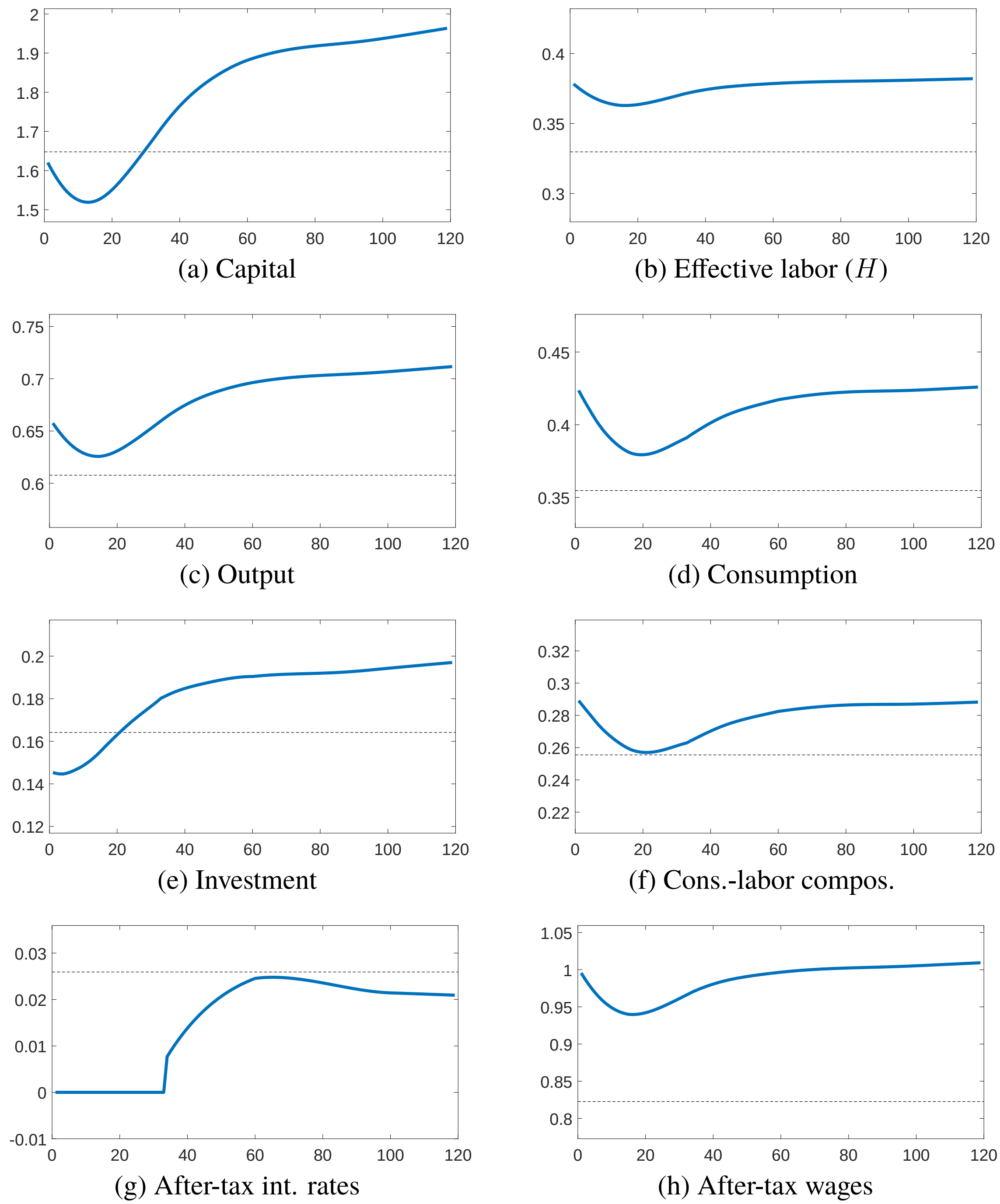
## C.3 Economy 4

**Proof of Proposition 6.** Equation (5.3) can be established for all $t \geq 1$, by substituting (8.19) into (8.20). The existence of a $t^*$ such that $\eta_t > 0$, for $t < t^* - 1$ and $\eta_t = 0$, for all $t \geq t^* - 1$, follows from a very similar logic to the one used in the proof of Proposition 4, which we suppress here[48]. Hence, for $t \geq t^*$ we can obtain $\tau_t^n$ by using (5.1), which follows from (8.19) with $\eta_t = 1$. For the same time period $\tau_t^k$ can then be found by using (5.3). Now, having $\tau_{t^*}^n$ we can use the fact that $\tau_t^k = 1$ and (5.3) moving backwards to obtain $\tau_t^n$ for $t < t^*$. ■

[48] With

$$P_{t^*} \equiv -\sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\Phi\Theta_{\tau-1} + \Psi\sigma}{\Upsilon_{\tau-1}\sigma}, \quad \text{and} \quad M_{t^*} \equiv \sum_{\tau=1}^{t^*} \frac{1}{\beta^\tau} \frac{\left(\prod_{t=\tau}^{t^*} R_t^* - 1\right)\Phi\Theta_{\tau-1} + \left(\frac{\Theta_{\tau-1}}{\Theta_{t^*}} \prod_{t=\tau}^{t^*} R_t^* - 1\right)\Psi\sigma}{\Upsilon_{\tau-1}\sigma}$$

# D Figures




(a) Capital
(b) Effective labor ($H$)
(c) Output
(d) Consumption
(e) Investment
(f) Cons.-labor compos.
(g) After-tax int. rates
(h) After-tax wages

Figure 7: Aggregates: Benchmark

Notes: Dashed line: initial stationary equilibrium; Solid line: optimal transition.