Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks*

Sebastian Dyrda
University of Toronto

Marcelo Pedroni
University of Amsterdam

First version: November 15, 2014
This version: November 7, 2018

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 with a new welfare-decomposition method. Optimal levels of capital and labor income taxes are roughly consistent with the prevailing ones in the US—in the long run for a utilitarian planner, and from the start for a planner that disregards equality concerns. High initial capital income taxes are an effective way to provide redistribution and are used in proportion to the planner's degree of inequality aversion. Optimal debt dynamics is substantially affected by the planner's degree of inequality aversion. The welfare function is relatively flat with respect to movements in long-run fiscal instruments. Ignoring transition or the dynamics of taxes over time can be severely misleading.

Keywords: Optimal Taxation; Heterogeneous 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, Chris Phelan, José-Víctor Ríos-Rull, Cezar Santos, Pedro Teles, 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, University of Western Ontario, LuBraMacro, and Banco de Portugal. This work was carried out in part using computing resources at the University of Minnesota Supercomputing Institute.

How should governments conduct fiscal policy in the presence of 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 heterogeneous 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 has been relatively successful at matching some basic facts about inequality and income risk.[1] In this environment, agents face risk 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 lead to endogenous inequality. As a result, on top of the usual concern about not distorting agents decisions, a 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 focused mainly 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 which amounts to the typical Ramsey optimal taxation problem. This paper is the first to directly address this problem. We calibrate the initial steady state to replicate several aspects of the US economy, in particular the fiscal policy, the distribution of wealth, earnings and income, and statistical properties of the individual labor income process. The final steady state is, then, endogenously determined by the optimal path of fiscal policy. As is 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 most of 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.

Main findings. Our main findings can be summarized as follows:

1. Optimal levels of capital and labor income taxes are roughly consistent with the prevailing ones in the US—in the long run for a utilitarian planner and from the start for a planner that disregards equality concerns. High initial capital income taxes and higher

[1] See, for instance, Domeij and Heathcote (2004); Castañeda, Díaz-Giménez and Ríos-Rull (2003); Heathcote, Storesletten and Violante (2009); and Nardi and Fella (2017).

[2] See, for instance, Aiyagari and McGrattan (1998), Conesa, Kitao and Krueger (2009), and Nakajima (2010). Notable exceptions are Krueger and Ludwig (2016) and Bakis, Kaymak and Poschke (2015) which solve for the optimal once-and-for-all change in policy and account for transitory effects.

overall labor income taxes are an effective way to provide redistribution and are used in proportion to the planner's degree of inequality aversion.

2. Optimal government debt dynamics is driven by the planner's degree of inequality aversion—the utilitarian planner calls for asset accumulation, a planner that disregards inequality for debt accumulation. Debt dynamics have significant general equilibrium price effect which, in turn, quantitatively important effects on the provision of insurance and redistribution.

3. The welfare function is relatively flat with respect to movements in long-run fiscal instruments. Our quantitative results show that the focus on long-run Ramsey policies in the SIM model is misguided.[3]

4. Ignoring transition or the dynamics of taxes over time can be severely misleading. Our contribution here is to quantify the importance of the transition both for the optimal fiscal instruments and for the associated welfare gains.

The optimal policy. For a utilitarian planner, our benchmark, we find that optimal capital income taxes are front-loaded, hitting the imposed upper bound of 100 percent for 53 years before decreasing to 42 percent in the long run. Labor income taxes gradually increase from 28 percent towards a final level of 41 percent. In the initial stationary equilibrium the capital and labor income taxes are set to their US counterparts: 36 and 28 percent, respectively. The ratio of lump-sum transfers to output is roughly doubled to about 16 percent and the government initially accumulates assets only to then return to a level of debt-to-output of about 40 percent in the long run—over the optimal transition government assets reach a level close to 100 percent of GDP. Relative to keeping fiscal instruments at their initial levels, this leads to a welfare gain equivalent to a permanent 13.9 percent increase in consumption.

Why use distortive taxes? Labor and capital income taxes are distortive; however, they are used to provide insurance and redistribution. The only risk that agents face, in the SIM model, is with respect to their labor productivities.[4] By taxing labor income and rebating the extra revenue via lump-sum, the planner reduces the proportion of the agents'

[3] By long-run we mean when the economy reaches the final stationary equilibrium. Part of the reason welfare is flat with respect to long-run instruments is because the convergence to the final steady state takes a very long time: about 200 years in our main results.

[4] 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 share of labor income above 77 percent in the 2007 SCF.

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, in this way, provides redistribution. To demonstrate exactly how the optimal policy reacts to changes in risk 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.

Welfare decomposition. To disentangle the main forces behind the optimal policy, we introduce a new procedure to decompose welfare gains[5] into what comes from the reduction of distortions to agent's decisions, from redistribution (the reduction of ex-ante risk) and from insurance (the reduction of ex-post risk). Applying this decomposition to our main results we find that the average welfare gains of 13.9 percent associated with implementing the optimal policy: (i) −4.8 percent comes from an increase in distortions to agents' decisions; (ii) 16.8 percent comes from redistribution; and (iii) 2.4 percent comes from the extra insurance provided by the fiscal policy. The optimal policy implies an overall increase of capital and labor income taxes which distort agents' savings and labor supply decisions more, leading aggregate resources to be less efficiently allocated. On the other hand, rebating the revenue of the higher taxes via lump-sum transfers, especially the higher capital income taxes in the initial periods when the links with the individual ex-ante states are still strong, decreases the proportion of the agents' income associated with the ex-ante unequal income and leads to the redistributional gains. Finally, the same mechanism acts to lower the proportion of the agents' income that is risky ex post leading to the positive insurance effect.

The role of redistribution. We also use the welfare decomposition to consider a Ramsey planner that disregards equality concerns. The welfare gains are, then, equivalent to a permanent 3.4 percent increase in consumption; 1.2 percent comes from the reduction in distortions and 2.2 percent from the provision of insurance. The much lower welfare gains are consistent with the fact that most welfare gains in the benchmark come from redis-

[5]The procedure is based on welfare decomposition used in Floden (2001) and Benabou (2002), but modified to handle transitory effects.

tribution. Optimal long-run capital and labor income taxes are set to 45 and 29 percent, respectively. Moreover, capital income taxes do not hit the upper bound, and instead move immediately to a value close to the long-run level, which is indicative that this feature of the benchmark results is mainly driven by the redistributive motive of the utilitarian planner. Lump-sum transfers are front loaded and, thus, the government accumulates debt, up to more than 300 percent of GDP. This relaxes borrowing constraints and provides insurance via general equilibrium price effects. We also consider an experiment allowing the planner to expropriate asset positions in the first period. Surprisingly, we find that even though 99 percent of asset holds are expropriated, it is still optimal to keep high capital income taxes. They are still beneficial for insurance and redistribution purposes, and the downward distortions on capital accumulation are counteracted by the lower elasticity of savings (since agents are rebuilding their precautionary buffers) and the crowding in of capital that results from the immense revenue obtained via the levy. Finally, we study an alternative welfare function which allows us to directly control the degree of inequality aversion of the planner, the weight put on equality concerns. The more the planner "cares" about inequality the longer it keeps capital income taxes in the upper bound, the higher the labor income taxes, and the lower the long-run debt-to-output ratio.

The importance of transition. Disregarding transitory welfare effects in Ramsey problems can be severely misleading. To make this point, we compute the stationary fiscal policy that maximizes welfare in the final steady state and show that it is substantially different from the optimal ones, since it ignores transitory effects and the costs associated with accumulating capital and reducing government debt. Moreover, even when the transition is taken under consideration, but taxes are restricted to being constant over time, the results are misleading, especially if one is interested in the long-run values for the fiscal instruments. The optimal taxes in this case are close to the short-run values of the dynamic optimal ones with capital income taxes, for instance, being set to 96 percent.

The role of market incompleteness. To illustrate the role of market incompleteness and highlight why and how our results 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, we analytically characterize[6] 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

[6]We build on the methodology developed in Werning (2007).

capital income 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 income 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 with the provision of insurance. A contribution of this paper is to quantify the optimal long-run capital income taxes in the SIM model—which to our knowledge has not been done before—as well as to highlight the role of debt dynamics which is indeterminate in a complete-markets model.

Related Literature

Assuming the existence of a Ramsey steady state, Aiyagari (1995) provides a rationale for positive long-run capital income taxes in the SIM model.[7] A recent study by Chen, Chien and Yang (2018) revisits this result and shows that, with no bounds on government debt, a Ramsey steady state may not exist.[8] Relative to the environments in Aiyagari (1995) and Chen, Chien and Yang (2018) the Ramsey planner in this paper has access to lump-sum transfers as an additional instrument. Moreover, as in Aiyagari (1995) we require government debt to be bounded. We view the contribution of this paper as quantifying the gains from the optimal fiscal policy conditional on finitely many policy adjustments over time, which essentially ensures the existence of the long-run steady state of the economy. Importantly, our computational algorithm does not rely on properties of the Ramsey long-run steady state. Instead, we verify *ex post* and independently of the solution method whether theoretical properties of the Ramsey steady state hold. In particular, we find that in the long-run steady state the modified golden rule does (approximately) hold.

Gottardi, Kajii and Nakajima (2015) and Heathcote, Storesletten and Violante (2017) characterize analytically the optimal fiscal policy in stylized versions of the SIM model. Krueger and Ludwig (2018) do the same in an overlapping generations setup. Their ap-

[7] Assuming the existence of a Ramsey steady state, Aiyagari shows that the modified golden rule has to hold at the optimum. At that associated interest rate, aggregate savings of agents with precautionary motive for savings would grow without bounds. Hence, Aiyagari (1995) concludes that a positive capital income tax must be imposed. 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 income tax rebated via lump-sum.

[8] Chien and Wen (2017), in an analytically tractable version of the SIM model, show that the modified golden rule holds if and only if the government can issue a sufficient amount of debt to enable households to achieve full self-insurance.

proaches lead to elegant and insightful closed-form solutions. 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.

The set of papers that tackle the issue of characterizing the optimal transition in a quantitative framework with heterogeneity is limited.[9] Itskhoki and Moll (2018) study optimal dynamic development policies in an incomplete markets model where heterogeneous producers are subject to financial frictions. To solve for the optimal transition they adopt a similar approach to ours, parametetrizing the time paths of tax instrument using exponential function of time, we use the more flexible family of functions—cubic splines. Nuño and Thomas (2016) use a novel continuous-time technique to solve for optimal monetary policy, including optimal transition, in a version of the incomplete markets model with money. Ragot and Grand (2017) solve the Ramsey problem in the SIM model with aggregate technology shocks by truncating the histories of idiosyncratic shocks. Acikgoz, Hagedorn, Holter and Wang (2018) argue that the long-run fiscal policy, in the SIM model, can be characterized independently of initial conditions. They compute it and solve backwards for the optimal transition towards the initial steady state. In contrast, our paper solves the Ramsey problem forward, and highlights the quantitative importance of the short-run dynamics of the optimal policy.[10]

We also contribute to the literature studying the nexus between government debt and market incompleteness. In an influential paper 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. Their calibration procedure focuses on matching the properties of the labor income process. Röhrs and Winter (2017) 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 replicates wealth, income and earnings inequality as well as the statistical properties of the labor income, hence we capture the main forces determining the dynamics of government debt over optimal transition and in the long run. Bhandari, Evans, Golosov and Sargent (2017) investigate the role of government debt in an economy with heterogeneity and aggregate risk. They find that the introduction of ad hoc borrowing constraints is welfare improving since it allows the Ramsey planner to affect

[9] Though, there is a vast literature analyzing the optimal policy in the steady state - for instance Conesa, Kitao and Krueger (2009) in an overlapping generations SIM model - or the optimal policy in the long run including transitory effects - Krueger and Ludwig (2016) or Bakis, Kaymak and Poschke (2015).

[10] We discuss the relationship between our results and theirs in Section 4.6 and in the Online Appendix.

general equilibrium prices. A similar logic is relevant for understanding our benchmark results.

An extensive literature studies the Ramsey problem in complete-market economies with heterogeneity. The most well known result for the deterministic subset of these economies is due to Judd (1985) and Chamley (1986): capital income 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. However, this result has been challenged by Straub and Werning (2014) who show that it can indeed be optimal to tax capital in the long run. Chari, Nicolini and Teles (2018) remove the Ramsey planner's ability to expropriate initial capital holdings and show that long-run capital income taxes should again be set to zero. The Ramsey planner in this paper also wants to expropriate capital holdings, but for a different reason: not to mimic lump-sum taxes since those are available, but to provide redistribution. Our experiment considering a planner that does not want to provide redistribution is, in this sense, related to the one in Chari, Nicolini and Teles (2018), though we still find optimal long-run capital income taxes to be positive. 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. Greulich, Laczó and Marcet (2016) characterize and compute Pareto improving capital and labor income taxes in the same setup. In Section 7 we characterize analytically and provide quantitative results for the optimal fiscal policy with complete markets in our environment and link the results to these studies.

Dávila, Hong, Krusell and Ríos-Rull (2012) solve the problem of a planner in the SIM model 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 affects after tax prices directly to achieve the same goal. Section 1.4 contains a more detailed discussion of the relationship between our results and theirs.

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, the welfare decomposition procedure. Section 5 considers a planner that disregards equality concerns and a utilitarian planner that can choose to expropriate capital. Section 6 discusses the importance of considering transitory effects. Section 7 presents the build-up from the complete market economy results to our main results. Section 8 provides results for alternative welfare functions and calibrations and Section 9 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 government could finance all necessary revenue with this non-distortive 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 *risk economy*. Then, we shut down risk and introduce ex-ante wealth inequality; this is referred to as the *inequality economy*. Next we discuss the relationship with the infinite horizon problem.[11]

1.1 Risk 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 β. In the first period each agent is endowed with ω 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 π 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

[11]The Online Appendix discusses the case in which there is risk *and* inequality.

in period 2 the fraction of agents with e_L is π 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, τ^n and τ^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*
$$\max_{a, n_L, n_H} u(\omega - a, \bar{n}) + \beta \mathbb{E}\left[u(c_i, n_i)\right] \quad s.t. \; c_i = (1-\tau^n) w e_i n_i + \left(1 + \left(1-\tau^k\right) r\right) a + T;$$
2. $r = f_K(K, N)$, $w = f_N(K, N)$, *where* $N = \pi e_L n_L + (1-\pi) e_H n_H$;
3. *and,* $\tau^n w N + \tau^k r K = G + T$.

The Ramsey problem is to choose τ^n, τ^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.[12]

Assumption 1 *No income effects on labor supply and constant Frisch elasticity,* κ, *i.e.*
$$u_{cn} - u_{cc}\frac{u_n}{u_c} = 0, \quad \text{and} \quad \frac{u_{cc} u_n}{n\left(u_{cc} u_{nn} - u_{cn}^2\right)} = \kappa.$$

[12] In a similar two-period environment, Gottardi et al. (2014) characterize the solution to Ramsey problem without Assumption 1. However, they impose an alternative assumption about the sign of general equilibrium effects, which are satisfied under Assumption 1. 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).

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.[13] 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 risk economy, if u satisfies Assumption 1, 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 balances the budget.

Proof. See Appendix A.1. ■

The planner could choose to finance G using only 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 inelastically. In fact, 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) since a higher κ implies a lower τ^n. That is, the more responsive labor supply is to changes in labor income taxes the more distortive these taxes are and the planner chooses a lower labor income 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 τ^k so as to mitigate the distortion imposed by a positive τ^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 $e_L = 1 - \epsilon^{risk}/\pi$ and $e_H = 1 + \epsilon^{risk}/\left(1-\pi\right)$, so that ϵ^{risk} is a mean-preserving spread on the productivity levels. It is easy to see that if $\epsilon^{risk} = 0$ equation (1.1)

[13] 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 $\tau^k < (1+r)/r$.

implies that $\tau^n = 0$. The effect of an increase in ϵ^{risk} on the optimal τ^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 $\partial\nu/\partial\epsilon^{risk} > 0$ and $\partial\nu/\partial\tau^n < 0$, it follows that $\partial\tau^n/\partial\epsilon^{risk} > 0$, 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, ν, is in fact increasing in the level of risk ($\partial\nu/\partial\epsilon^{risk} > 0$) and decreasing in the labor income tax ($\partial\nu/\partial\tau^n < 0$).

1.2 Inequality economy

Consider the environment described above only without risk 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 ω can take two values: ω_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*

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

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 $W \equiv pU_L + (1-p)\,U_H$, known as the utilitarian welfare function. The following proposition follows.

Proposition 2 *In the inequality economy, if* u *satisfies Assumption 1 and has CARA or is GHH, as in equation (3.1), then the optimal tax system is such that* $\tau^n = 0$,

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

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* ψ *is the level of absolute risk aversion.*[14] T *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 income 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 τ^k. Hence, the higher this elasticity is the lower is the optimal level of τ^k, since providing redistribution becomes more costly. The $\tau^n = 0$ result is again associated with Assumption 1.

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

1.3 Relationship with Dávila, Hong, Krusell and Ríos-Rull (2012)

The results established in Dávila et al. (2012) have an interesting relationship to the ones we obtain in this paper. We use the last result to explain this relationship. Among other things, Dávila et al. (2012) show that the competitive equilibrium allocation in the SIM model is constrained inefficient. That is, the incomplete market structure itself induces outcomes that could be improved upon if consumers merely acted differently, that is if they used the same set of markets but departed from purely self-interested optimization. The constrained inefficiency results from a pecuniary externality. The savings and labor supply decisions of

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

the agents affect the wage and interest rates and, therefore, the risk and inequality in the economy. These effects are not internalized by the agents and inefficiency follows. Note that the planner's problem in their environment is significantly different from the Ramsey problem described here. There the planner affects allocations directly and prices indirectly, as a result redistribution and insurance can only occur via the manipulation of equilibrium prices. Whereas here the Ramsey planner affects (after tax) prices directly and allocations indirectly.

In a setting similar to the inequality economy just described above, for instance, Dávila et al. (2012) show that there is under accumulation of capital. A higher level of capital would decrease interest rates and increase wages, reducing inequality. A naive extrapolation of this logic would suggest that capital income taxes should be negative so as to encourage savings. This logic, however, does not take into account the more relevant direct effect of the tax system on after tax prices. Proposition 2 shows that the opposite is true: capital income taxes should be positive.

1.4 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 (ex-post) risk and inequality (ex-ante risk). In the infinite horizon version of the SIM model, however, these concepts are inevitably intertwined. The characterization of the optimal tax system, therefore, becomes considerably more complex. Labor income taxes affect not only the level of risk 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.

2 The Infinite-Horizon Model

Time is discrete and infinite, indexed by t. There is a continuum of agents with standard preferences $\mathbb{E}_0\left[\sum_t \beta^t u\left(c_t, n_t\right)\right]$ 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 over time governed by

Γ, a transition matrix.[15] 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 taxes $\{\tau_t^n\}_{t=0}^{\infty}$, (positive) capital income taxes $\{\tau_t^k\}_{t=0}^{\infty}$, 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, τ^c, is a parameter.[16] Let $\{\lambda_t\}$ be a sequence of probability measures over the Borel sets $\mathcal{S}$ of S with λ_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 $f\left(\cdot\right)$ that uses capital, K_t, and efficient units of labor, N_t, to produce output each period ($f\left(\cdot\right)$ denotes output net of depreciation, δ 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

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

[15] 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 Γ is such that there exists a unique $p^* = \lim_{t \to \infty} p_t$. We normalize $\sum_i p_i^* e_i = 1$.

[16] It is not without loss of generality that we do not allow the planner to choose τ^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 τ^c as a parameter for calibration purposes.

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, $\{\tau_0^k, \tau_0^n, T_0\}$ an initial distribution λ_0 and a policy $\pi \equiv \{\tau_t^k, \tau_t^n, T_t\}_{t=1}^{\infty}$, 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 π, $c_t(a,e)$, $n_t(a,e)$, 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 λ_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, λ_0, and $\{\tau_0^k, \tau_0^n, T_0\}$, for every policy π,* ***equilibrium allocation rules*** *$X(\pi)$ and* ***equilibrium price rules*** *$P(\pi)$ are such that π, $X(\pi)$, $P(\pi)$ and corresponding $\{v_t\}_{t=0}^{\infty}$ and $\{\lambda_t\}_{t=1}^{\infty}$ constitute a competitive equilibrium.*

Definition 5 *Given K_0, B_0, λ_0, and $\{\tau_0^k, \tau_0^n, T_0\}$, and a welfare function $W(\pi)$, the **Ramsey problem** is to $\max_{\pi \in \Pi} W(\pi)$ such that $X(\pi)$ and $P(\pi)$ are equilibrium allocation and price rules, and Π is the set of policies $\pi = \{\tau_t^k, \tau_t^n, T_t\}_{t=1}^{\infty}$ for which an equilibrium exists.*[17]

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 λ_0. The planner's objective is, thus, given by

$$W(\pi) = \int_S \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(a, e|\pi\right), n_t\left(a, e|\pi\right)\right) d\lambda_0.$$

We consider alternative welfare functions in Sections 5.1 and 8.1.

2.2 Solution method

We solve the Ramsey problem defined above numerically. Given an initial stationary equilibrium, for any policy $\pi \in \Pi$ we can compute the transition to a new stationary equilibrium consistent with that policy, as long as the taxes become constant at some point, and evaluate welfare $W(\pi)$. We then search for the policy $\pi = \{\tau_t^k, \tau_t^n, T_t\}_{t=1}^{\infty}$ maximizing welfare $W(\pi)$. This is, however, a daunting task since it involves searching in the space of infinite sequences. In order to make it computationally feasible we approximate the space of infinite policy sequences Π with a space of sequences, Π_A, that can be identified by a finite number of nodes.

In Section 7 we show that in complete markets economies optimal capital income taxes should be front-loaded. Hence, in defining the set Π_A we take this under consideration. That is, we allow capital income taxes to hit the imposed upper bound of 100 percent for the first t^* periods, where a model period is equivalent to one year. Importantly, t^* is a choice variable and is allowed to be zero, so the fact that the solution displays capital income taxes 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 $\{\tau_t^n, T_t\}_{t=1}^{\infty}$ follow splines with nodes set at exogenously selected periods. 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 nodes: t^*, $\tau_{t^*+1}^k$, τ_{75}^k, τ_{100}^k, τ_1^n, τ_{15}^n, τ_{30}^n, τ_{45}^k, τ_{60}^k, τ_{100}^k, T_1, T_{15}, T_{30}, T_{45}, and T_{60}. The last node for lump-sum, T_{100}, is determined endogenously in order for government debt to be bounded and is, therefore, not a choice variable. In

[17] In particular, note that the government debt path associated with a policy in Π must be bounded.

the intermediate periods the paths follow a cubic spline function and after the final period, 100, they become constant at the last level. The choice of the periods 1, 15, 45, 60, and 100, were placed at the same distance from each other except for the last ones which are supposed to capture the long run levels. The choice of nodes for $\left\{\tau_t^k\right\}_{t=t^*}^{100}$ are a result of the fact that, for experiments with less nodes, the optimal t^* was always close to 50. In the Online Appendix we include details about the calculation of T_{100} and figures that compare the optimal fiscal policy computed with 2, 3, 6, 8, 10, 13 and, finally, 15 variables. The welfare gains associated with each of these solutions are displayed in Table 1—the gains are computed as permanent percent increases in consumption. We view the fact that the changes in welfare gains from 13 to 15 nodes is small, around 0.01 percent, and that the optimal taxes in the two experiments are close to one another as evidence that we have allowed for enough nodes.

Table 1: Change in welfare at optimum from adding nodes

Nodes	$2 \rightarrow 3$	$3 \rightarrow 6$	$6 \rightarrow 8$	$8 \rightarrow 10$	$10 \rightarrow 13$	$13 \rightarrow 15$
Welfare	0.54	0.27	0.11	0.26	0.03	0.01

Solving the problem described above is a particularly hard computational task. Effectively we are maximizing $W(\pi)$ such that $\pi \in \Pi_A$. 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 design a numerical algorithm for global optimization, based on insights from Guvenen (2011), Kan and Timmer (1987a) and Kan and Timmer (1987b), and apply it to solve the Ramsey problem. Our algorithm is parallelized for multiple cores and a detailed description of it is contained in the Online Appendix. Here, we present a heuristic overview.

The algorithm can be divided into two main stages: a global and a local one. In the global stage we randomly draw a very large number of policies from the approximated domain Π_R and compute the transition between the exogenously given initial stationary equilibrium and a final stationary equilibrium that is policy dependent. Then, we compute welfare $W(\pi)$ for each of those policies and select the ones that yield the highest levels of welfare. These selected policies are then clustered, i.e. similar policies in terms of welfare are placed in the same cluster. Next, in the local stage we run for each cluster a derivative free optimizer based on an algorithm designed by Powell (2009). The sequence of global and local searches is repeated until the number of local minima found and the expected number of local minima in our problem, determined by a Bayesian rule, are sufficiently close. Then we pick the global optimum from the set of local optima. The main experiment, with 15 nodes, was conducted with the use of 576 cores at the Minnesota Supercomputing Institute and took approximately 120 hours.

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 2 summarizes our 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[18] and from the 2007 Survey of Consumer Finances (SCF).

Table 2: Benchmark Model Economy: Target Statistics and Parameters

Statistic	**Target**	**Model**	**Parameter**	**Value**
Preferences and Technology				
Intertemporal elast. of substitution	0.50	0.50	σ	2.000*
Frisch elasticity	0.72	0.72	κ	0.720*
Average hours worked	0.30	0.30	χ	3.905
Capital to output	2.72	2.72	β	0.948
Capital income share	0.38	0.38	α	0.380*
Investment to output	0.27	0.27	δ	0.100
Borrowing Constraint				
% of hhs with wealth < 0	18.6	19.3	$\underline{a}/Y$	-0.025
Fiscal Policy				
Capital income tax (%)	36.0	36.0	τ^k	0.360*
Labor income tax (%)	28.0	28.0	τ^n	0.280*
Consumption tax (%)	5.0	5.0	τ^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

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

[18] 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 σ is the coefficient of relative risk aversion, κ is the Frisch elasticity of labor supply and χ 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), Chetty et al. (2012) and Cesarini et al. (2017) for details.[19]

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, κ, 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 χ is chosen[20] so that average hours worked equals 0.3 (the associated average effective labor level, N, is 0.44). To pin down the discount factor, β, we target a capital to output ratio of 2.72, and the depreciation rate, δ, is set to match an investment to output ratio of 27 percent.[21] The aggregate technology is given by a Cobb-Douglas production function $Y = K^{\alpha}N^{1-\alpha}$ with capital share equal to α, which 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.

[19] 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.

[20] 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.

[21] 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. 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 Röhrs and Winter (2017)). 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.

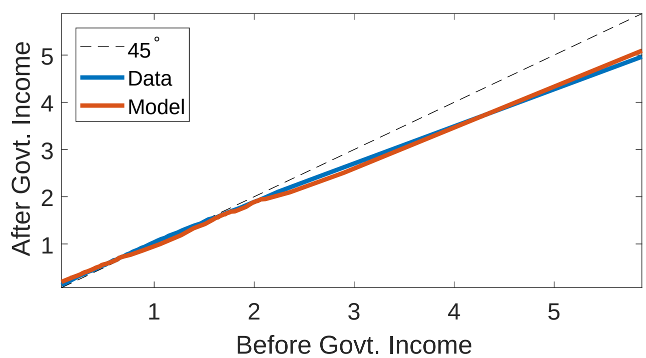


Figure 1: Income tax schedule

Note: The data was generously supplied by Heathcote et al. (2017) 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 stochastic process for individual labor productivity levels, e, is calibrated to match statistical properties of the labor income process and the distributions of wealth, earnings and income. We model it as a sum of a persistent component e_P with Markov matrix Γ_P and a transitory component e_T with probability vector P_T.[22] There are 4 persistent and 6 transitory productivity levels. Since we normalize the average productivity to one and probabilities must also add up to one, we are left with 26 parameters to choose.

It is common to use the discretization procedures introduced by Tauchen (1986) or Rouwenhorst (1995) when calibrating the Markov process for productivities. These methods have limited ability to represent higher order moments of the labor income process

[22]In the notation of the model, $e = e_T + e_P$ and $\Gamma = \Gamma_P \otimes \text{diag}(P_T)$.

Table 3: Benchmark Model Economy: Target Statistics and Parameters

Model Parameters

Persistent Shock

$$\Gamma_P = \begin{bmatrix} 0.96 & 0.04 & 0.00 & 0.00 \\ 0.08 & 0.91 & 0.01 & 0.00 \\ 0.01 & 0.00 & 0.98 & 0.01 \\ 0.09 & 0.01 & 0.01 & 0.89 \end{bmatrix} \quad e_P = \begin{bmatrix} 0.48 \\ 0.88 \\ 1.80 \\ 7.14 \end{bmatrix}$$

Transitory Shock

$$P_T = \begin{bmatrix} 0.05 \\ 0.04 \\ 0.12 \\ 0.59 \\ 0.10 \\ 0.11 \end{bmatrix} \quad e_T = \begin{bmatrix} -0.23 \\ -0.10 \\ -0.06 \\ 0.01 \\ 0.18 \\ 0.21 \end{bmatrix}$$

Statistics	**Target**	**Model**	**Target**	**Model**	**Target**	**Model**
Moment of Distribution	**Wealth**		**Earnings**		**Income**	
Gini index	0.82	0.84	0.64	0.61	0.57	0.55
% in bottom 5%	−0.2	−0.0	−0.1	0.5	0.2	0.9
% in 1st quintile	−0.2	−0.1	−0.1	3.4	2.8	4.6
% in 2nd quintile	1.1	0.3	4.2	4.1	6.7	5.8
% in 3th quintile	4.5	2.1	11.7	8.3	11.3	9.6
% in 4th quintile	11.2	10.8	20.8	19.7	18.3	22.6
% in 5th quintile	83.4	86.9	63.5	64.5	60.9	57.5
% in top 5%	60.3	58.4	35.3	34.5	36.9	31.2
Statistical Properties of Labor Income Process						
Variance of 1-year diff.	0.26	0.27				
Skewness of 1-year diff.	−1.07	−0.75				
Kurtosis of 1-year diff.	14.93	14.58				
Variance of 5-year diff.	0.61	0.64				
Skewness of 5-year diff.	−1.25	−0.82				
Kurtosis of 5-year diff.	9.51	10.19				
Autocorrelation	0.88	0.88				

such as its skewness and kurtosis. We do not impose the restrictions associated with those methods which allows us to have more targets.

To match the wealth, earnings and income distributions we follow Castañeda et al. (2003) and target the Gini coefficients, the shares owned by every quintile, plus the bottom and top 5 percent in the 2007 Survey of Consumer Finances, as reported by Díaz-Giménez et al.

(2011). Similarly to Domeij and Heathcote (2004) we also target properties of the individual labor income[23] estimated by Guvenen et al. (2015). Specifically, we target the variance, skewness and kurtosis of labor income growth in 1 and 5 years, and the autocorrelation of the annual labor income.[24] Table 3 summarizes the parameter values and targeted moments, notice that the model is over identified with 31 targets for the 26 parameters.

3.5 Model performance

Table 4 presents an important dimension along which our model is consistent with the data: income sources over the quintiles of income. The composition of income,[25] especially of the consumption-poor agents, plays an important role in determining 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.

Table 4: Income sources by quintiles of income

Quintile	US Data			Model		
	Labor	Asset	Transfer	Labor	Asset	Transfer
1st	38.4	-1.9	63.5	57.0	0.7	42.3
2nd	66.4	2.5	31.1	60.8	5.9	33.3
3rd	78.6	2.7	18.7	73.9	6.1	20.0
4th	85.4	4.0	10.6	71.0	20.4	8.5
5th	77.5	18.2	4.3	78.4	18.3	3.3
All	77.3	12.2	10.4	74.3	16.1	9.6

Note: 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, based on a summary by Díaz-Giménez et al. (2011).

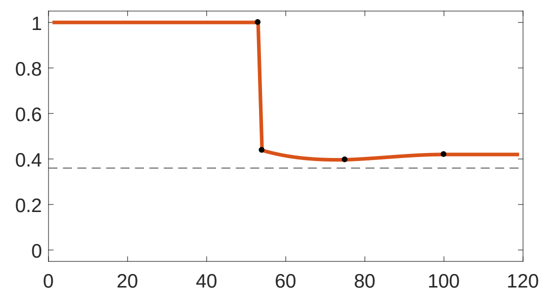
[23] Although in our model there is no difference between the two concepts, in dealing with the data we follow Díaz-Giménez et al. (2011) and call labor income wages and salaries of all kinds, and earnings the sum of labor income plus a fraction of business income.

[24] We compute these moments in closed form directly from the Markov matrix relying on insights from Civale et al. (2016).

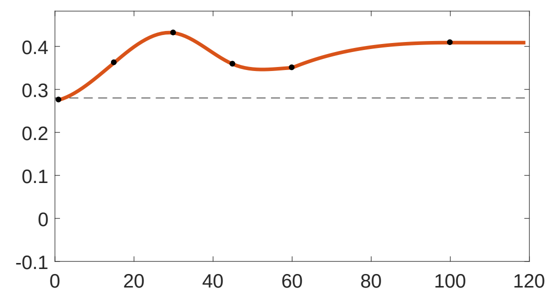
[25] Dávila et al. (2012) demonstrate that the composition of income, especially of poor agents, is a crucial determinant of the optimal policy.

4 Main Results

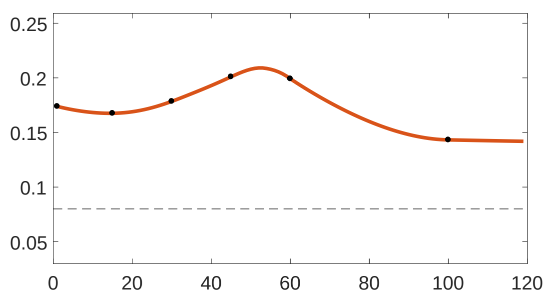
The optimal paths for the fiscal policy instruments are presented in Figure 2. Capital income taxes are front-loaded hitting the upper bound for 53 years, and then decrease to 42 percent in the long run. Labor income taxes initially remain close to the initial value of 28 percent, then increase towards a final value of 40 percent. The ratio of lump-sum transfers to output is more than doubled from the initial value of 8 percent to 17 percent initially, eventually settling on about 14 percent. The government accumulates assets in the initial periods of high capital income taxes reaching a level of debt-to-output of about -100 percent, which then converges to a final level of 40 percent. Relative to keeping fiscal instruments at their initial levels, this leads to a welfare gain equivalent to a permanent 13.9 percent increase in consumption. This section is devoted to explaining the economics behind these results.



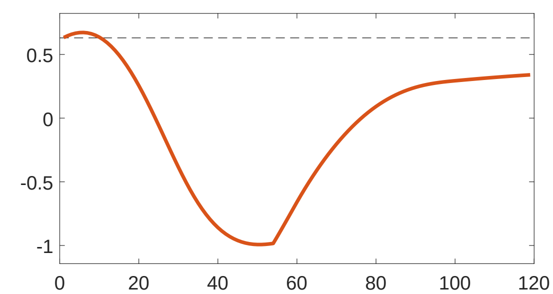
(a) Capital income tax



(b) Labor income tax



(c) Lump-sum-to-output



(d) Debt-to-output

Figure 2: Optimal Fiscal Policy: Benchmark

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

4.1 Aggregates

Looking only at the aggregates it seems hard to justify the optimal policy. This is because the welfare gains associated with the policy come from the implied redistribution and extra insurance which require higher taxes and, therefore, are likely to be detrimental to aggregate movements, a point that we clarify bellow in Section 4.3. The aggregates associated with the implementation of the optimal policy are shown in Figure 15. At the end of the transition capital is reduced by 28 percent, and labor by 17 percent.

Since there are no wealth effects on labor supply, the reduction of aggregate labor is easy to understand it is a result of the higher labor income taxes and the lower level of capital. The movement in capital has more forces at play. Besides the higher overall capital income taxes and lower aggregate labor which reduces the marginal product of capital, another force that acts to reduce the capital level is the reduction of the precautionary savings due to the fact that the optimal policy implies a less risky after-tax labor income. Note, however, that even if capital income taxes were set to 100 percent forever, there would still be a precautionary motive to save. Moreover, the fact that the government accumulates assets over time, especially during the years with capital income taxes at 100 percent, crowds in capital which also limits its reduction - an effect we explain in more detail below in Section 4.7.

The lower levels of capital and labor lead to lower levels of output and, therefore, aggregate consumption, which decrease by 24 and 26 percent respectively, over the transition. The concomitant reduction in average consumption and labor has ambiguous effects on the welfare of the average agent. Hence, we also plot in Figure 15f 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 17 percent since, besides the reduction in labor, consumption levels increase due to the initial reduction in savings. It then decreases over time, eventually reaching a level that is 20 percent lower than in the initial steady state.

4.2 Distributional Effects

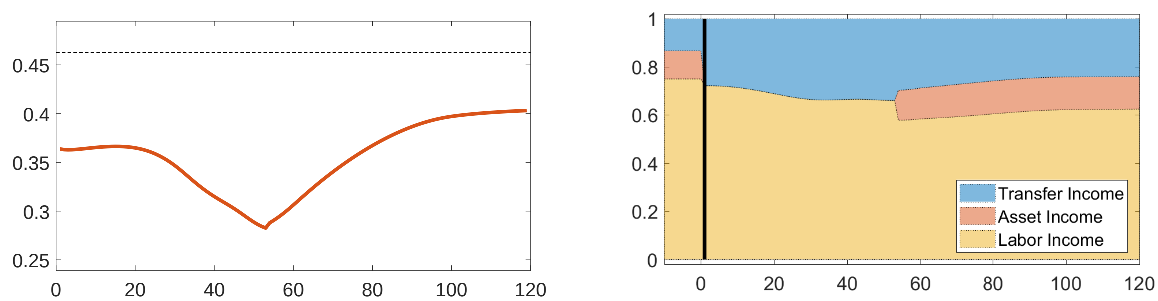
As discussed above, movements in the aggregates do not provide a full picture of what results from the implementation of the optimal fiscal policy. It is also important to understand its effects on inequality and on the risk faced by the agents. Figure 3a plots the evolution of the Gini index for consumption-labor composite. Note that, on impact the Gini is significantly reduced and that this reduction is mostly maintained over the transition.

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 only by the agents with lower asset positions and productivity levels - see Table 5.

Table 5: Proportion in favor of reform by earnings and wealth quintiles

Quintile	1st	2nd	3rd	4th	5th	All
Earnings	88.1	88.1	88.0	62.3	0.0	65.3
Wealth	99.8	99.1	89.6	27.8	9.1	65.3

Figure 3b displays the evolution of the shares of labor, capital and transfer income out of total income. It is important to notice that the share of labor income is significantly reduced under the optimal policy and replaced mostly by transfer income. Since all the risk faced by agents in the SIM model is associated with their labor income, it turns out that they face less risk after the policy is implemented, which is also welfare improving. The next sections will quantify more precisely the importance of these effects.



(a) Cons.-Labor Comp. Gini (b) Income shares

Figure 3: Inequality measures

Note: (a) Dashed line: initial stationary equilibrium; Solid line: optimal transition; (b): From top to bottom the areas represent the shares of transfer, (after-tax) asset and labor income; before time 0 the areas represent the shares in the initial stationary equilibrium.

4.3 Sources of welfare improvement

Here we present a result that is particularly helpful for understanding the properties of the optimal fiscal policy. First, let $v(x_t) \equiv u(c_t, n_t)$ where x_t is the individual consumption-labor composite, the term “inside” the utility function in equation (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, $U(\{X_t\}) \equiv \sum_{t=0}^{\infty} \beta^t v(X_t)$, 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 riskiness of 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 risk faced by the agents. Finally, it increases 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 what follows we define these components precisely and present propositions to support their usefulness.

Average welfare gain. Let $v(x_t) \equiv u(c_t, n_t)$ where u is defined in (3.1) and consider a policy reform. Denote by $x_t^R(a_0, e^t)$ 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 $x_t^{NR}(a_0, e^t)$ be the equilibrium path in case there is no reform. The average welfare gain, Δ, that results from implementing the reform is defined as the constant (over time and across agents) percentage increase to $x_t^{NR}(a_0, e^t)$ that equalizes the utilitarian welfare to the value associated with the reform, that is,

$$\int \mathbb{E}_0 \left[U\left((1+\Delta) \left\{ x_t^{NR}\left(a_0, e^t\right) \right\} \right) \right] d\lambda_0\left(a_0, e_0\right) = \int \mathbb{E}_0 \left[U\left(\left\{ x_t^{R}\left(a_0, e^t\right) \right\} \right) \right] d\lambda_0\left(a_0, e_0\right), \tag{4.2}$$

where λ_0 is the initial distribution over states (a_0, e_0) and

$$U\left(\left\{x_t\left(a_0, e^t\right)\right\}\right) \equiv \sum_{t=0}^{\infty} \beta^t v(x_t(a_0, e^t)) = \sum_{t=0}^{\infty} \beta^t u\left(c_t\left(a_0, e^t\right), n_t\left(a_0, e^t\right)\right).$$

Components of welfare. Let the aggregate level of x_t at each t be

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

Then, the level effect, Δ_L, is given by

$$U\left((1+\Delta_L)\left\{X_t^{NR}\right\}\right) = U\left(\left\{X_t^R\right\}\right). \tag{4.3}$$

Let $\{\bar{x}_t^j(a_0, e_0)\}$ denote the sequence of individual consumption-labor certainty equivalents,

$$U\left(\left\{\bar{x}_t^j(a_0,e_0)\right\}\right) = \mathbb{E}_0\left[U\left(\left\{x_t^j\left(a_0,e^t\right)\right\}\right)\right], \quad \text{for } j = R, NR. \tag{4.4}$$

Note that the path $\{\bar{x}_t^j(a_0, e_0)\}$ is not fully determined by this condition. We therefore, impose that for every initial state (a_0, e_0), the individual certainty equivalent is proportional to the path of an agent who always has the productivity level e_0 and has assets a_0 in period 0, that is[26]

$$\bar{x}_t^j(a_0,e_0) = \eta^j(a_0,e_0)\, x_t^j(a_0,e_0), \quad \text{for } j = R, NR, \tag{4.5}$$

where $\eta^j(a_0, e_0)$ denotes the degree of proportionality. The usefulness of this condition will become apparent briefly. Next, let $\bar{X}_t^j$ be the aggregate consumption-labor certainty equivalent,

$$\bar{X}_t^j = \int \bar{x}_t^j(a_0,e_0)\, d\lambda_0(a_0,e_0), \quad \text{for } j = R, NR. \tag{4.6}$$

The insurance effect, Δ_I, is defined by

$$1+\Delta_I \equiv \frac{1-p_{risk}^R}{1-p_{risk}^{NR}}, \text{ where } \quad U\left(\left(1-p_{risk}^j\right)\left\{X_t^j\right\}\right) \equiv U\left(\left\{\bar{X}_t^j\right\}\right), \tag{4.7}$$

and the redistribution effect, Δ_R, by

$$1+\Delta_R \equiv \frac{1-p_{ineq}^R}{1-p_{ineq}^{NR}}, \text{ where } \quad U\left(\left(1-p_{ineq}^j\right)\left\{\bar{X}_t^j\right\}\right) \equiv \int U\left(\left\{\bar{x}_t^j(a_0,e_0)\right\}\right) d\lambda_0(a_0,e_0). \tag{4.8}$$

The terms p_{risk} and p_{ineq} are the costs of risk and inequality in the economies with or without reform. These definitions have the following useful properties.

Proposition 3 *If there is no risk in the economy, the cost of risk, p_{risk}, is zero, and if there is no inequality, the cost of inequality, p_{ineq}, is equal to zero.*

Proof. See Appendix B. ■

The definition of the individual certainty equivalents in equation (4.5) is crucial to estab-

[26] This way of defining individual certainty equivalents is the main difference between our welfare decomposition procedure and the one used in Floden (2001) and Benabou (2002); it is the key to establishing Proposition 3.

lishing this result. When computing this decomposition in a stationary equilibrium it is common to set them to a constant. When the transition is take into account, however, such a choice leads to non-zero risk costs even when there is no risk in the economy. Besides this being an oxymoron, it would mean that the magnitude of the insurance effect would not be completely due to changes in the amount of risk the agents are exposed to.[27]

Welfare decomposition. The following proposition establishes that it is possible to decompose the average welfare gains into the components described above.[28]

Proposition 4 *If preferences are GHH, then*

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

Proof. See Appendix B. ■

Note that none of the elements of the decomposition are defined residually, hence this is indeed a decomposition and not a definition. The results of applying this decomposition for our main results are in Table 6.

Table 6: Welfare decomposition

	Δ	Δ_L	Δ_I	Δ_R
Benchmark	13.9	−4.8	2.4	16.8
Fixed Capital Income Tax	3.2	−2.5	2.1	3.6
Fixed Labor Income Tax	12.7	−1.8	0.4	14.3
Fixed Debt-to-Output	12.0	−7.3	2.5	17.9
Fixed Prices	17.0	−1.3	1.5	16.8

Fixed Instruments. Besides the benchmark economy, we also include in Table 6 the decomposition for four other economies, in which we keep each fiscal instrument at their initial level. Other instruments follow their optimal paths found in the main experiment and we add a constant to the path o lump-sum so that the government budget constraint is satisfied. Most of the welfare gains, in the benchmark experiment, come from the redistribution channel, the insurance channel contributes 2.4 percentage points, and the level effect has a negative contribution of −4.8 percent.

[27]We thank Piero Gottardi for pointing this out.

[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 Δ_L, Δ_I, and Δ_R accordingly.

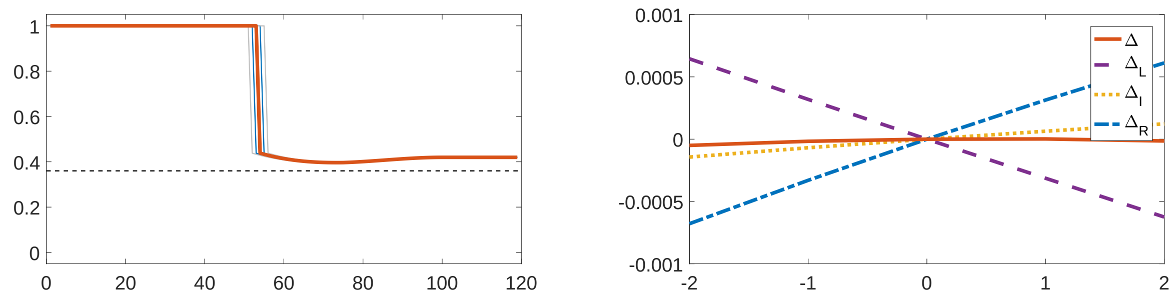
It is clear from the results the path of capital income taxes is behind most of the redistributional welfare gains while labor income taxes are behind most of the insurance gains. On the other hand keeping these taxes at their initial lower levels leads to less welfare losses through the level effect. These results are fully in line with Propositions 1 and 2. Keeping debt-to-output at its initial level, exacerbates the damage to the level effect caused by the long period high capital income taxes at the upper bound, since the government no longer has an effective instrument to crowd-in private capital through general equilibrium prices. On the other hand, the gains via redistribution increase since lump-sum transfers are forced to rise more during this initial period. Finally, to understand the contribution of general equilibrium price movements, we consider an out-of-equilibrium experiment in which prices are kept at their initial values while all fiscal instruments move as in the benchmark. Since aggregate output decreases in the benchmark results, with fixed prices agents will have relatively more income which explains the smaller reduction in the level effect. In the next sections we elaborate on the reasons behind these results.

4.4 Variations around the optimal taxes

In this section we vary the tax levels around the optimal values and calculate the welfare decomposition at each step in order to better understand the main determinants of the optimal values. For every experiment, the entire path of lump-sum taxes was shifted up or down in order to balance the government's intertemporal budget constraint.

Number of years of capital income taxes in the upper bound. The optimal path of capital income taxes features 53 years of taxes at the imposed upper bound of 100 percent, which we denote by $t^* = 53$. Figure 4 shows what happens to the components of welfare if capital income taxes are kept at the upper bound for more or less periods. The effect on insurance is of second order, and, in line with the result in Proposition 2, the relevant trade-off is between the extra redistribution associated with a higher t^* versus the negative level effect due to the extra amount of distortion. These two effects, however, almost exactly offset each other, leading to a relatively flat average welfare function which is consistent with the findings in Section 4.5.

Long-run capital income taxes. Here we vary all the nodes for capital income taxes after t^* by the same amount. The level of capital income taxes at these nodes is allowed to increase or decrease up to 5 percent relative to the optimal one. The results are displayed in Figure 5 and the welfare decomposition numbers are both comparable to the ones for changes in t^*. That is, again the relevant trade-off is between the redistribution and the



(a) Capital income tax (b) Welfare decomposition

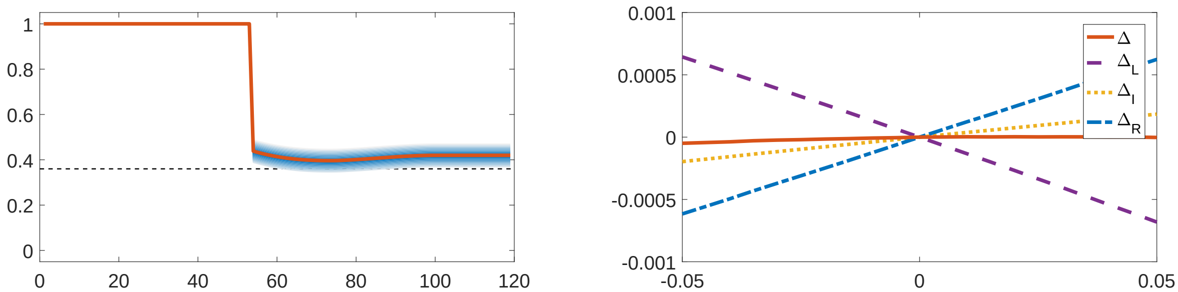
Figure 4: Variation on t^*

Note (a): Dashed line: initial stationary equilibrium; Solid line: optimal transition and variation upon it; (b): the x-axis represents the movement in t^* for the optimum minus 2 to plus 2, and the lines plot the difference between each welfare measure and its value at the optimal.

level effects. This is somewhat surprising since these tax changes only affect capital income taxes after 53 years of transition and one might expect that the dependence on agents' initial condition would have mostly dissipated by then. The fact that this is not the case speaks to how persistent this dependence actually is.[29] For the logic in Chamley (2001) and Acikgoz et al. (2018)—that far enough in the long-run the dependence on agents' initial conditions fully dissipates so that only the insurance and level effects would be relevant—to become relevant one would need to consider changes in capital income taxes further in the future than we consider here. Movements far enough in the future would indeed only affect the insurance effect and have no effect on ex-ante redistribution. On a related note, in Section 5.1 we show that the insurance effect by itself can rationalize levels of capital income taxes very similar to the long-run levels seen here. Finally, take notice of the range of values for the welfare decomposition in Figure 5. The low order of magnitude of these numbers is indicative of the relevance for welfare of the long-run capital income taxes in this model but also, the fact that these figures display monotonic and well behaved curves testifies to the precision of our solution.

Labor income taxes. Here we change the average level of labor income taxes up and down by 5 percent, leading to the results in Figure 6. First note the by comparing the welfare numbers with the ones in Figures 4 and 5, the effect of changes in labor income taxes are an order of magnitude higher than the changes to capital income taxes considered above. Besides this quantitative difference, the main qualitative difference is that the insurance effect plays a comparable role to the redistribution effect in determining the optimal level of

[29] Even focusing only on the productivity states, we have that $\max_{i,j} |(\Gamma^{100})_{i,j} - \pi_j| \approx 0.1$, that is, the probability of being in state j in period 100 and having started in state i is still significantly different from the probability of being in state j in according to the stationary distribution, denoted by π.

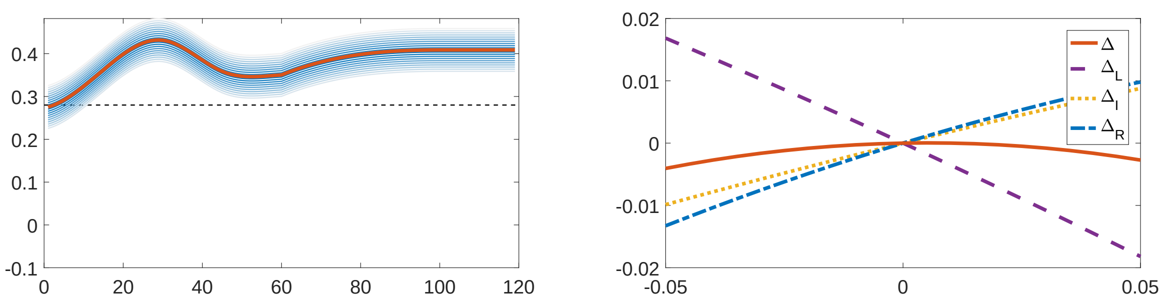


(a) Capital income tax (b) Welfare decomposition

Figure 5: Variation on long-run capital income taxes

Note (a): Dashed line: initial stationary equilibrium; Solid line: optimal transition and variation upon it; (b): the x-axis represents the movement in all nodes of capital income taxes (after t^*) from the optimum minus 5 to plus 5 percent, and the lines plot the difference between each welfare measure and its value at the optimal.

labor income taxes. Hence, though labor income taxes do have important effects on ex-ante risk, the mechanism highlighted in Proposition 1 plays an important role here. That is, a higher labor income tax which is rebated via lump-sum (exactly the experiment here), effectively reduces the labor income risk that agents are exposed to.



(a) Labor income tax (b) Welfare decomposition

Figure 6: Variation on labor income taxes

Note (a): Dashed line: initial stationary equilibrium; Solid line: optimal transition and variation upon it; (b): the x-axis represents the movement in t^* for the optimum minus 5 to plus 5 percent, and the lines plot the difference between each welfare measure and its value at the optimal.

4.5 Flatness of Welfare Function

Our numerical procedure for solving the Ramsey problem allows us to probe some aspects of the solution that are hard to investigate analytically. In the process of optimizing, the solver tries several different paths for the fiscal instrument in its search for the optimal one. We save all these paths and in Figure 7 present the ones that lead to the highest welfare.

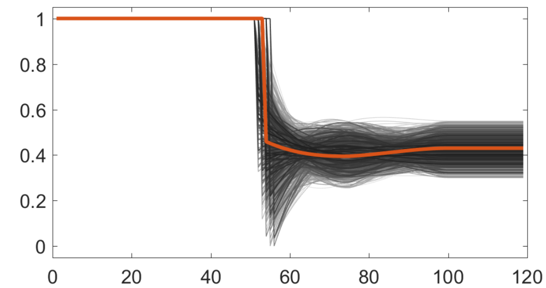
All paths plotted in this figure are associated with a welfare gain higher than 13.85 percent, less than 0.01 percent away from the optimal path - the darkness of a path indicating higher welfare.

First note that the short-run dynamics is much more important, with no path deviating substantially from the optimal in the first 40 years. Though this is obviously a result of the fact that the planner discounts the future (at the same rate as agents), it does undermine to some extent discussions about the precise values of the long-run optimal taxes in this model; it does not have a significant effect on welfare. Further, note that the long-run levels of capital income taxes and debt-to-output are the ones that display the most amount of variation. Since the optimal taxes resolve trade-offs between redistribution, insurance and the amount of distortions to agents decisions, it is not surprising that the welfare function would be relatively flat around the optimum. An increase in capital income taxes, for instance, would improve welfare by providing more redistribution and insurance but reduce it since it distorts the agents' saving decision. These effects exactly offset each other at the optimum, so "small enough" movements in the fiscal instruments do not matter much for welfare. Figure 7 illustrates how small is small enough.

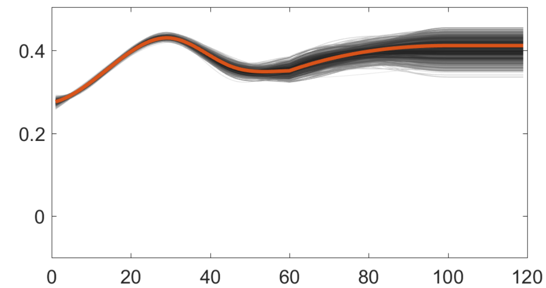
4.6 Long-Run Optimality Conditions

Aiyagari (1995) analyses the optimal long-run capital income taxes in an environment similar to the one we are working with.[30] He argues that, since there is no aggregate risk, 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 income taxes. This logic also implies that the modified golden rule should hold in the long run; our numerical results imply exactly that. Figure 8 displays $\beta(1 + f_K(K, N))$ for our benchmark results (solid line) and for an once-and-for-all policy change experiment, discussed in more detail in Section 6.2 (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.

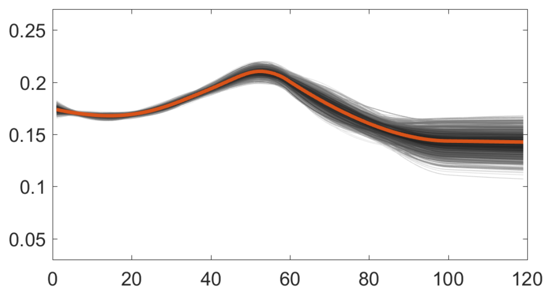
[30] The home production assumption in Aiyagari (1995) is equivalent to our assumption that preferences are GHH. The differences are that in his environment the planner does not have lump-sum taxes as an instrument, but chooses the level of government expenditure every period (which enters separately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



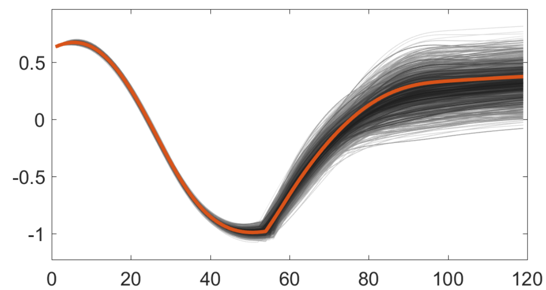
(a) Capital income tax



(b) Labor income tax



(c) Lump-sum-to-output



(d) Debt-to-output

Figure 7: Flatness of Welfare Function

Note: Around the optimal taxes, we plot all (1563) paths that generate welfare gains within 0.01 percent of the optimal path; the darkness of a path indicates higher welfare.

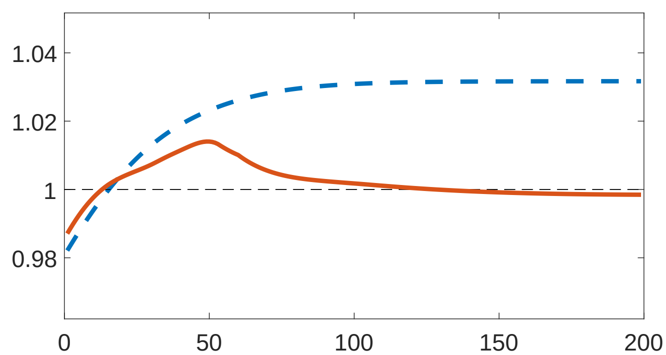


Figure 8: $\beta(1 + f_K(K, N))$

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

Recently Acikgoz et al. (2018) have made advancements towards obtaining a better characterization of the long-run optimal tax system in environments very similar to ours. They argue that the long-run optimal tax system is independent of initial conditions and of the transition towards it, and show that three long-run optimality conditions must be satisfied—the modified golden rule and two additional ones. They propose an algorithm that allows for the computation of the optimal long-run tax system. We have applied this algorithm to our economy and have found consistent long-run results.[31] The results in Acikgoz et al. (2018) are obtained under the assumption that a Ramsey steady state exists, that is, the optimal Ramsey policy is such that the policy itself and all equilibrium variables converge to a steady state. Chen et al. (2018) show, in a similar environment, that a Ramsey steady state should not exist.

One possible explanation for these seemingly contradicting results is that, in Chen et al. (2018), lump-sum transfers are not an instrument to the Ramsey planner. Also, both Acikgoz et al. (2018) and this paper assume, as a no-Ponzi condition, that government debt must be bounded. Chen et al. (2018) do not impose such a constraint, in which case it could be that the planner chooses a path of debt that grows without bounds—a weaker no-Ponzi condition, on the growth rate of debt, can still be satisfied. Either way, in this paper, we restrict government debt to be bounded and our solution method only approximates the solution to the Ramsey problem if a Ramsey steady state exists. In that case, the long-run optimality conditions in Acikgoz et al. (2018) should be satisfied and it is reassuring that they are. Our solution method, however, does not make use of these conditions in any way; they are verified ex post. Our results, moreover, indicate that these long-run properties have very small implications for welfare since they are only relevant very far in the future and that a focus on the short-run dynamics is more relevant.

4.7 The role of government debt

In the absence of borrowing constraints, an increase in government debt, financed by an appropriate change in the timing of lump-sum transfers, is innocuous. In response, agents simply adjust their savings one-to-one and the Ricardian equivalence holds.[32] 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.[33]

[31] 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.

[32] Since we are referring to the effect of changes in the timming of lump-sum transfers financed by debt, the violation of Ricardian equivalence associated with proportional taxation, as in Barsky et al. (1986), is not an issue.

[33] See Aiyagari and McGrattan (1998) and Röhrs and Winter (2017) for an extensive discussion of this issue.

In the benchmark experiment this mechanism is mostly used by the Ramsey planner to smooth out the capital path over time. The reduction in debt in the initial periods crowds in capital, which counterbalances the effect of the high capital income taxes. However, movements in government debt can also have important general equilibrium price effects. A lower level of government debt, for instance, leads to a higher capital level, 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. These effects are going to play an important role in what follows.

5 The role of redistribution

In this section we consider two experiments that aim to clarify exactly what features of the benchmark results are driven solely by the redistributive motive of the planner. Besides providing additional insights about those results we find these experiments intrinsically interesting as well.

5.1 Maximizing efficiency

The redistributive motive of the planner plays a central role in our benchmark results. This preference for reducing inequality is associated with the particular welfare weights of the utilitarian welfare function. Here, we use the welfare decomposition, explained in Section 4.3, and consider the problem of a Ramsey planner that maximizes the level and insurance effects of the welfare decomposition, $(1 + \Delta_L)(1 + \Delta_I)$. In the equality versus efficiency trade-off, such a planner places no weight on equality, focusing only on the reduction of distortions and ex-post risk.[34] Figure 9 presents the results in comparison with the benchmark results. Relative to the initial stationary equilibrium, the welfare gains associated with the policy are equivalent to a permanent 3.4 percent increase in consumption, 1.2 percent coming from the reduction in distortions and 2.2 percent from the extra insurance.

Labor and capital income taxes. Relative to the benchmark experiment, labor income taxes are about 10 percent lower throughout the transition. As indicated in Section 4.4, higher labor income taxes are beneficial both for insurance and redistributive motives, so

[34] In Section 8.1 we consider different levels of inequality aversion.

it makes sense that removing one of these motives from consideration leads to a lower labor income tax. Long-run capital income taxes are remarkably similar to the benchmark one. In the short run, however, capital income taxes are not kept at the upper bound for any periods. In the benchmark results, capital income taxes are front loaded because the utilitarian planner wants to provide redistribution and, since the dependence of agent's individual states on their initial values dissipates over time, the earlier redistribution takes place the better. Since lump-sum transfers in the benchmark results could be reduced in every period, the usual explanation for the front loading of capital income taxes—that it mimics lump-sum taxes as it distorts less the agents' decisions the earlier the taxes are imposed—does not hold. However, just like in a representative agent economy without lump-sum, what the planner really would like to do is to confiscate the initial asset holdings of the agents, though here the goal is to redistribute them.[35]

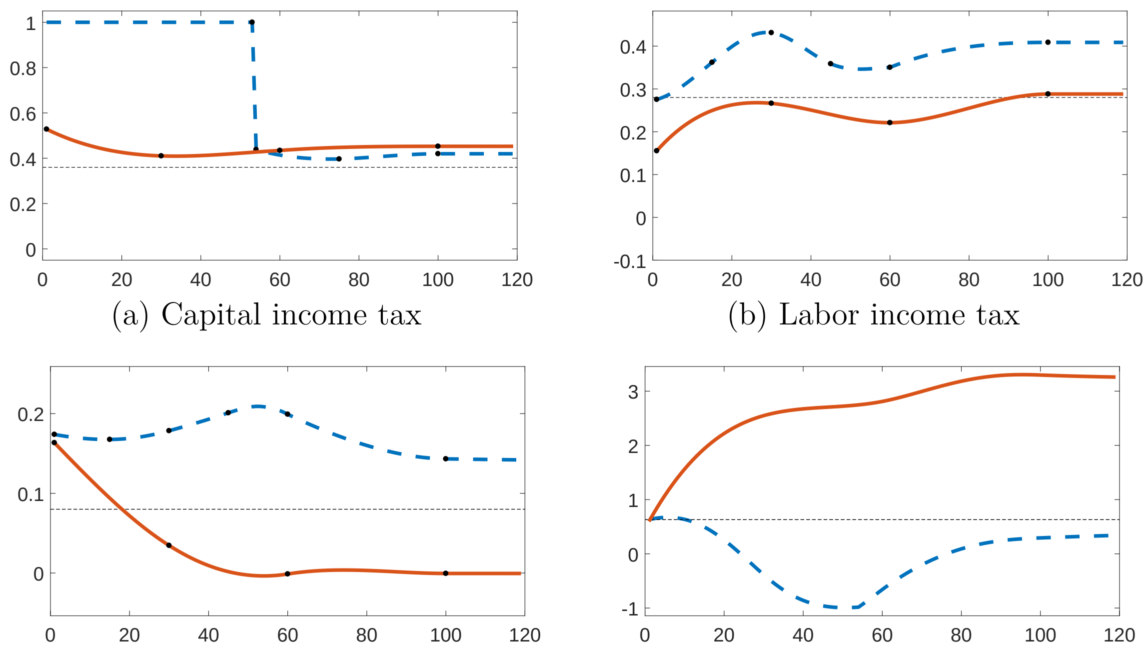
Lump-sum transfers and debt. Since average (over-time) taxes on capital and labor income are lower than in the benchmark, lump-sum transfers must, on average, also be lower. To understand the path of lump-sum over time it is relevant to notice that, absent borrowing constraints, the agents would be indifferent about its timing. Since the agents face borrowing constraints, it is, therefore, optimal to front load lump-sum transfers as much as possible. This, however, has the negative side effect of increasing government debt, which crowds out capital. Incidentally, the reason why lump-sum transfers are not front loaded in the benchmark experiment is because this crowding out in combination with the capital income taxes at 100 percent would lead to a fast and substantial reduction in the capital level. Here, debt-to-output increases steadily towards a final level of 322 percent, but capital levels are still higher than in the benchmark experiment throughout the transition as a result of the lower overall capital and labor income taxes.[36] The general equilibrium price effects associated with the crowing out of capital—the reduction in wages and increase in interest rates—is also responsible for the insurance gains that follow from implementing this policy.

Share of constrained agents. In the benchmark result, the share of constrained agents increases as a side effect of redistribution via capital income taxation, since it effectively compresses asset positions towards the constraint. Having a larger share of agents close to the borrowing constraint is also beneficial in the sense that it gives the planner more

[35]The experiment of considering a planner that ignores redistributive concerns is, therefore, akin to the experiment in Chari et al. (2018) where they restrict policies from reducing the value of initial wealth in utility terms which effectively removes the planner's motive to expropriate initial asset holdings.

[36]The Online Appendix contains the figures for the aggregates of this experiment and the next where we allow for an initial capital levy.

power to affect capital via changes in debt; the Ricardian equivalence holds to a lower degree—this effect is highlighted by Bhandari et al. (2017). On the other hand, being close to the constraint is detrimental to the insurance effect since agents are less able to absorb negative shocks. Since the planner here cares relatively more about insurance, it actually brings to zero the share of agents that are borrowing-constrained by front loading lump-sum transfers; in the benchmark results, 9 percent of agents are borrowing-constrained in the long run.



(a) Capital income tax

(b) Labor income tax

(c) Lump-sum-to-output

(d) Debt-to-output

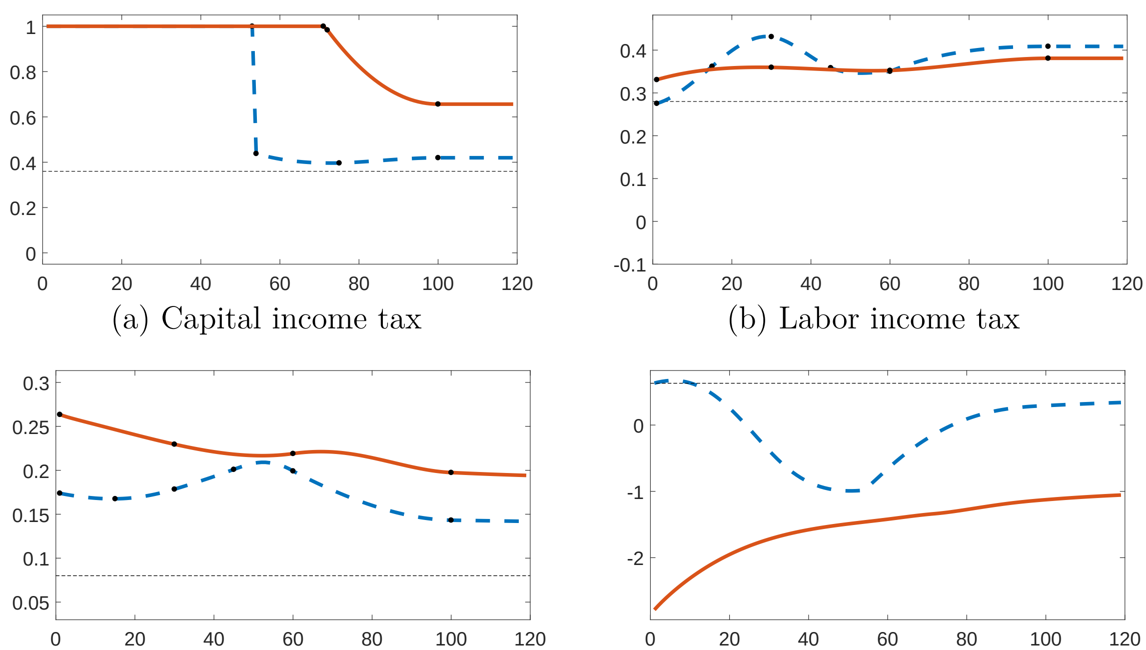
Figure 9: Optimal Fiscal Policy: Maximizing Efficiency

Note: Thin dashed line: initial stationary equilibrium; Solid line: path that maximizes efficiency optimal transition; Thick dashed line: path that maximizes the utilitarian welfare function (Benchmark Results).

5.2 Initial capital levy

Since the utilitarian planner wants to front load capital income taxes, we conduct another experiment in which we allow the planner to also change capital income taxes in period 0. We find that the planner then chooses to expropriate 99 percent of the initial asset holding. Surprisingly, however, this does not lead to lower capital income taxes in the future periods, on the contrary, capital income taxes are kept in the upper bound of 100 percent for longer (71 years) and reach a higher long-run level than in the benchmark experiment.

With the capital levy, high productivity agents, who have an incentive to accumulate a precautionary buffer of savings, have virtually all their assets expropriated in period 0. These agents immediately begin to rebuild the asset positions and these efforts are not significantly diminished by the high capital income taxes. That is, the savings decision of the agents who save the most in this economy becomes even less elastic. Moreover, on impact the government obtains a lot of revenue so their asset position goes immediately from a debt-to-output level of 63 percent to -279 percent. As a result, capital is crowded in and the downward distortions to capital accumulation associated with capital income taxes are less relevant. On the other hand, capital income taxes are still beneficial to provide insurance and further redistribution going forward. Importantly, even though capital income taxes are overall significantly higher relative to the benchmark, the equilibrium capital stock is still higher throughout the transition. The results are shown in Figure 10. The welfare gains are equivalent to a permanent 32.4 percent increase in consumption, -2.8 percent coming from the level effect, 1.9 percent from the insurance and 33.6 percent from redistribution.



(a) Capital income tax

(b) Labor income tax

(c) Lump-sum-to-output

(d) Debt-to-output

Figure 10: Optimal Fiscal Policy: Levy on Initial Capital Income

Note: Thin dashed line: initial stationary equilibrium; Solid line: path that maximizes the utilitarian welfare function allowing for capital income taxes to move in period 0 (though the tax level at $t = 0$ is not plotted since it is equal to $2,570$ percent) ; Thick dashed line: benchmark results.

6 Transitory effects

In this section we quantify the importance of transitory effects. 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 Table 7.

Table 7: Final Stationary Equilibrium: transitory effects

	τ^k	τ^n	T/Y	B/Y	K/Y	Δ	Δ_L	Δ_I	Δ_R
Initial equilibrium	36.0	28.0	8.0	63.0	2.71	–	–	–	–
Stat. equil.	–	43.4	16.1	−375.2	4.20	41.3	6.7	0.1	32.3
Stat. equil. no debt	32.6	27.7	7.6	63.0	2.76	0.1	0.9	−0.1	−0.7
Constant policy	96.1	34.9	20.6	−95.5	2.02	12.7	−4.5	1.8	15.9
Benchmark	42.0	40.9	13.9	40.0	2.48	13.9	−4.8	2.4	16.8

Note: All values, except for K/Y, are in percentage points.

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 −375.2 percent. At this level the amount of capital that is crowded in is close to the golden 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 income taxes which is why we do not display that number in Table 7. The average welfare gains associated with this policy are 41.3 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 transitory effects, it is as if the economy jumped immediately to a new steady state with a much higher capital stock and in which the government has a large amount of assets without incurring in the costs associated with accumulating it.

This result contrasts with the one in Aiyagari and McGrattan (1998). They run a similar experiment with some important differences: in their model the only tax available to the planner is a total income tax, and their calibration strategy for the labor income process focuses on matching the auto-correlation and variance of labor income without targeting distributional moments. They find that the government, even though it could costlessly choose any level of debt-to-output, chooses a level very close to the actual level in the US data at the time, around 67 percent. In fact, they show that the welfare function is relativelly flat with respect to the choice of debt-to-output. Röhrs and Winter (2017) 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 (they also allow for different tax instruments which they show also affects the result). The mechanism described in Section 4.7 is important to understanding these results: a higher debt level crowds out capital, which increases interest rates and reduces wages. Therefore, a higher debt level generates (i) a positive insurance effect, and (ii) negative level and redistribution effects. In Aiyagari and McGrattan (1998) the former dominates the latter, yet here and in Röhrs and Winter (2017) the opposite occurs. In particular, matching the actual level of wealth inequality makes redistribution a higher priority for the utilitarian planner. Figure 11 presents the welfare gains and decomposition for an experiment in which we allow only for total income taxes that balance the budged, as in Aiyagari and McGrattan (1998), and vary the level of debt-to-output.

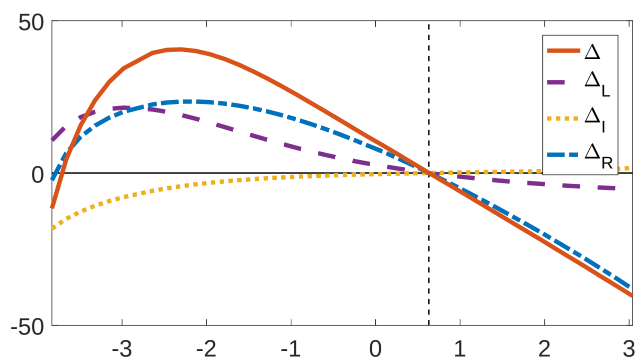
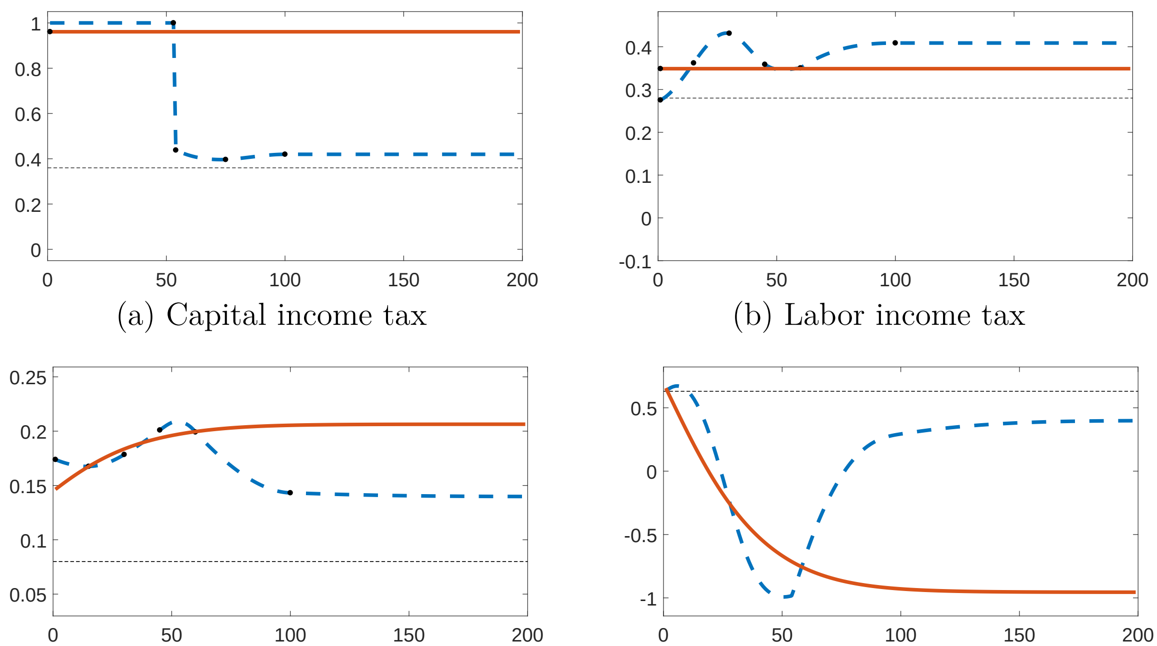


Figure 11: Welfare decomposition versus debt-to-gdp in steady state

Note: The variable in the x-axis is the debt-to-gdp in steady state; the thin dashed vertical line marks the level of debt-to-gdp in the initial stationary equilibrium, 63%, versus which the welfare changes are calculated.

An alternative experiment, which is closer to the one studied by Conesa et al. (2009), is to restrict the level of debt-to-output to remain at its initial level and choose only the other fiscal instruments. When this is the case, the planner chooses fiscal instruments at levels very close to the initial ones. As a result, the policy leads to small welfare gains of only 0.1 percent relative to the initial steady state. Interestingly, this result for the fiscal instruments is analogous to the finding in Aiyagari and McGrattan (1998) about the level of debt-to-output. However, implementing the steady-state-maximizing policy and accounting for its transitory effects would actually lead to a welfare loss equivalent to a 0.7 percent permanent reduction in consumption.

6.2 Transition with constant policy



(a) Capital income tax

(b) Labor income tax

(c) Lump-sum-to-output

(d) Debt-to-output

Figure 12: Optimal Fiscal Policy: Constant Policy

Notes: Thin dashed line: initial stationary equilibrium; Solid line: path that maximizes efficiency optimal transition; Thick dashed line: Benchmark Results.

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. Since the government cannot change taxes over time and the welfare function puts higher weight on the short run, the optimal taxes under this restriction are close to the ones in the short run of the optimal dynamic taxes in the benchmark experiment (see Figure 12).[37] The long-run levels of the fiscal instruments, however, are significantly

[37]Figures with the corresponding aggregates are presented in the Online Appendix.

different. Long-run capital income taxes and debt-to-output are especially different, since they vary more over the transition. Hence, if one is interested in the long-run properties of the fiscal instruments, it is important to allow them to vary over time. In particular, as we noticed above in Section 4.6, whereas the modified golden rule (approximately) holds for the benchmark policy, it does not hold under this restriction (see Figure 8). Finally, note the restriction to constant policies leads an average welfare loss of 1.2 percent relative to the optimal dynamic policy.

7 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, Γ, 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.[38] Figure 13 contains the numerical results obtained using the same method used for the benchmark results together with some of the analytical equations derived bellow.

7.1 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.

[38] In order to keep the amount of labor income inequality comparable with the benchmark calibration we rescale the productivity levels so as to keep the variance of the present value of labor income the same.

This amounts to setting $\tau_t^k = 0$ and $\tau_t^n = -\tau^c$ for all $t \geq 1$. Since consumption taxes are exogenously set to a constant level, zero capital income taxes leave savings decisions undistorted and labor income taxes set 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.

7.2 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)[39] 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 5 *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.[40] 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, Greulich, Laczó and Marcet (2016) obtain analogous results. In Section 8.1 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 income 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 income taxes can be used to provide redistribution. This fact together with the fact that capital income taxes are zero in the

[39] 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.

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

long run determine the optimal path for capital income taxes.[41] Capital income taxes are positive and front-loaded, hitting the upper bound in the initial periods and subsequently being set 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 to achieve redistribution.

In order to have a sense of the magnitude 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,[42] hence, we can choose any one we want. To keep the results comparable we choose the initial stationary distribution from the benchmark experiment.[43]

7.3 Economy 3: heterogeneity in productivity levels

It turns out that the Ramsey policies for this economy are a bit more complex. Let Φ, Ψ, and Ω^n be constants, defined in Appendix C, and define

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

Proposition 6 *Assuming capital income taxes are bounded only by the positivity of gross interest rates, the optimal labor income tax, τ_t^n, can be written as a function of Θ_t given by*

$$\tau_t^n(\Theta_t) = \frac{(1+\tau^c)\,\Psi\Theta_t}{\Phi\Theta_t + \Psi(\sigma+\Theta_t)} - \tau^c, \quad \text{for } t \geq 1, \tag{7.1}$$

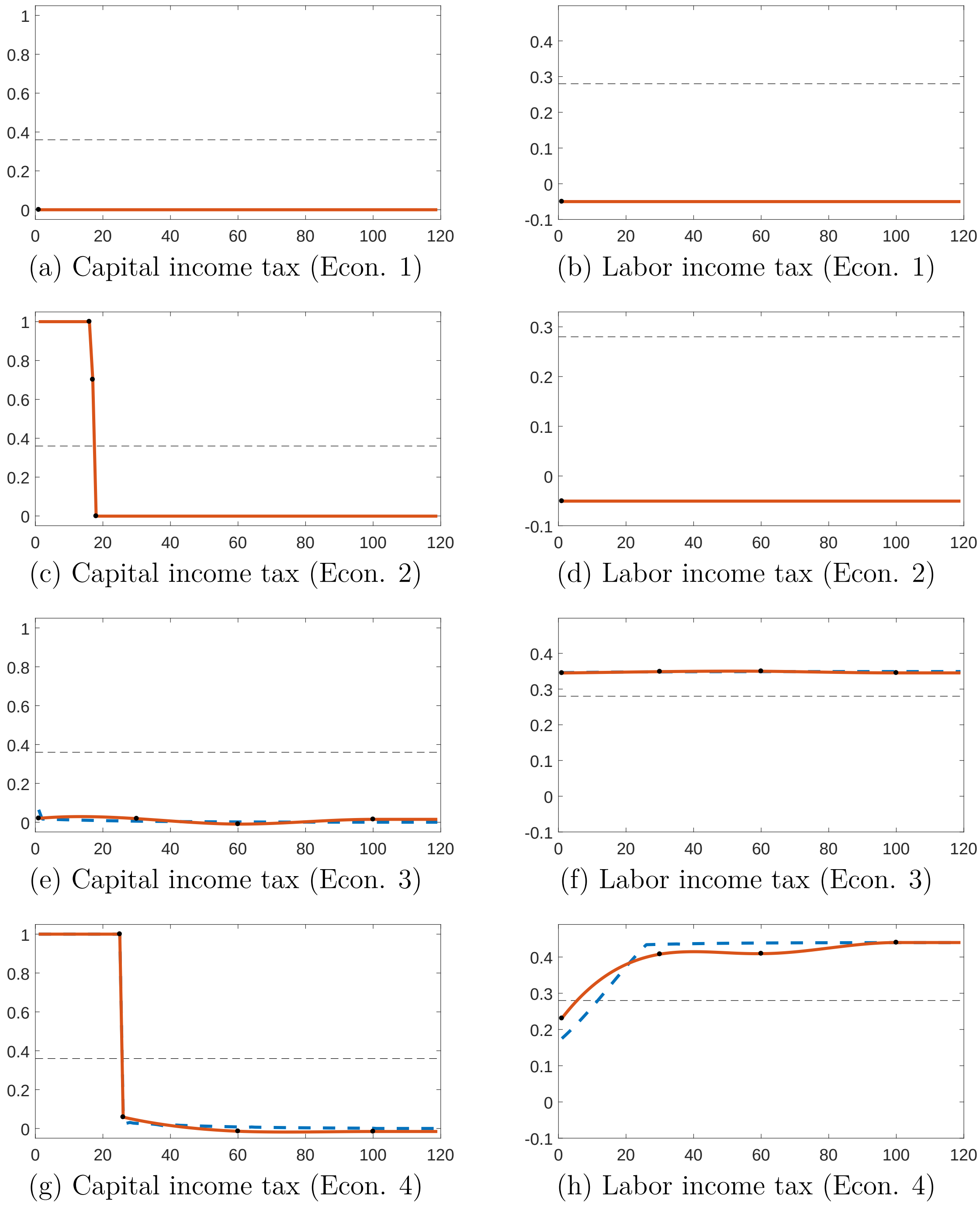
with sensitivity

$$\Theta_t \frac{d\tau_t^n(\Theta_t)}{d\Theta_t} = \frac{\sigma\left(\tau_t^n(\Theta_t)+\tau^c\right)^2}{(1+\tau^c)\,\Theta_t}. \tag{7.2}$$

[41] Straub and Werning (2014) show that optimal long-run capital income 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.

[42] For the preferences chosen above, consumption is linear in 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.

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



(a) Capital income tax (Econ. 1) (b) Labor income tax (Econ. 1)
(c) Capital income tax (Econ. 2) (d) Labor income tax (Econ. 2)
(e) Capital income tax (Econ. 3) (f) Labor income tax (Econ. 3)
(g) Capital income tax (Econ. 4) (h) Labor income tax (Econ. 4)

Figure 13: Optimal Taxes: Complete Market Economies

Notes: Thin dashed line: initial taxes; Solid line: optimal taxes calculated using the same procedure used in the Benchmark experiment; Thick dashed line: optimal taxes calculated by using the proposition equations.

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

$$\frac{1+(1-\tau^k_{t+1})r_{t+1}}{1+r_{t+1}} = \frac{\tau^n_t+\tau^c}{\tau^n_{t+1}+\tau^c}\frac{1-\tau^n_{t+1}}{1-\tau^n_t}, \quad \textit{for } t \geq 1. \tag{7.3}$$

Proof. See Appendix C.2. ∎

Since labor income is unequal, there is a redistributive reason to tax it. Optimal labor income taxes are not constant over time since they depend on Θ_t. If they were constant, however, equation (7.3) would imply $\tau^k_t = 0$ for all $t \geq 2$. Thus, capital income taxes will fluctuate around zero to the extent that labor income taxes vary over time. We disregard the upper bound on capital income taxes, $\tau^k_{t+1} \leq 1$, because it would complicate the result even further and in a non-interesting way. It could be that the bound is violated if the variation of Θ_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 τ^n_t to Θ_t we plug the initial stationary equilibrium numbers ($\tau^n = 0.28$, $\tau^c = 0.05$, $\sigma = 2$, and $\Theta \approx 2$) into equation (7.2). This implies a sensitivity of 0.1, i.e. a 1 percent increase in Θ_t changes the tax rate by 0.1 of a percentage point, from 0.28 to 0.2797.[44] This fact, together with the relative stability of Θ_t over time, implies that the optimal labor income taxes are virtually constant and capital income 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 income taxes vary over time, they distort the savings decision, and capital income 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.

For this economy and the next, Figure 13 presents the optimal taxes calculated two ways: using the same procedure as in the benchmark experiment, and using the equations from the propositions. We view the fact that the two are very similar as a validation of the procedure used to obtain the benchmark results.

[44] We can also calculate the path of Θ_t, which we displayed in a figure in the Online Appendix.

7.4 Economy 4: heterogeneity in initial assets and productivity levels

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

Proposition 7 *There exists a finite integer $t^* \geq 1$ such that the optimal tax system is given by $\tau_t^k = 1$ for $1 \leq t < t^*$, τ_t^k follows equation (7.3) for $t > t^*$; τ_t^n evolves according to equation (7.3) for $1 \leq t < t^*$; and τ_t^n is determined by equation (7.1) for all $t \geq t^*$.*

Proof. See Appendix C.3. ■

Optimal capital income taxes are very similar to those of Economy 2, and for the same reasons. Labor income taxes are determined by the same equation as in Economy 3 for $t \geq t^*$. In initial period, $1 \leq t < t^*$, while capital income taxes are at the upper bound, $R_t = 1 < R_t^*$ and, therefore, equation (7.3) implies that labor income taxes should be increasing. Lump-sum transfers are higher than in Economies 2 and 3 since they are used to redistribute the capital *and* labor income tax revenue.

Importantly, the optimal labor income taxes are quantitatively similar to the benchmark results and its pattern over time, and displays a similar qualitative feature, i.e. while capital income taxes are at the upper bound, labor income taxes are increasing. This pattern follows immediately from equation (7.3) by setting $\tau_{t+1}^k = 1$. The high capital income tax level distorts savings downwards, so having labor income taxes increase over time "undoes" this distortion to some extent as it front-loads (after-tax) labor income which increases savings. The only important qualitative difference between these results and the benchmark ones are that here, capital income taxes are set to zero in the long run.

8 Robustness

Figure 14 shows that the solution with 4 nodes (t^*,$\tau_{t^*+1}^k$,τ_1^n, and T_1) produces a reasonable approximation for the benchmark solution, at least with respect to its basic features, leading to welfare gains of 13.2 percent relative to 13.9 percent in the benchmark results. In this section we use this approximation to explore to evaluate the robustness of the results with respect to changes in the planner's degree of inequality aversion, the labor-supply and intertemporal elasticities, and the introduction of preference shocks such that labor supply is independent of the productivity level.

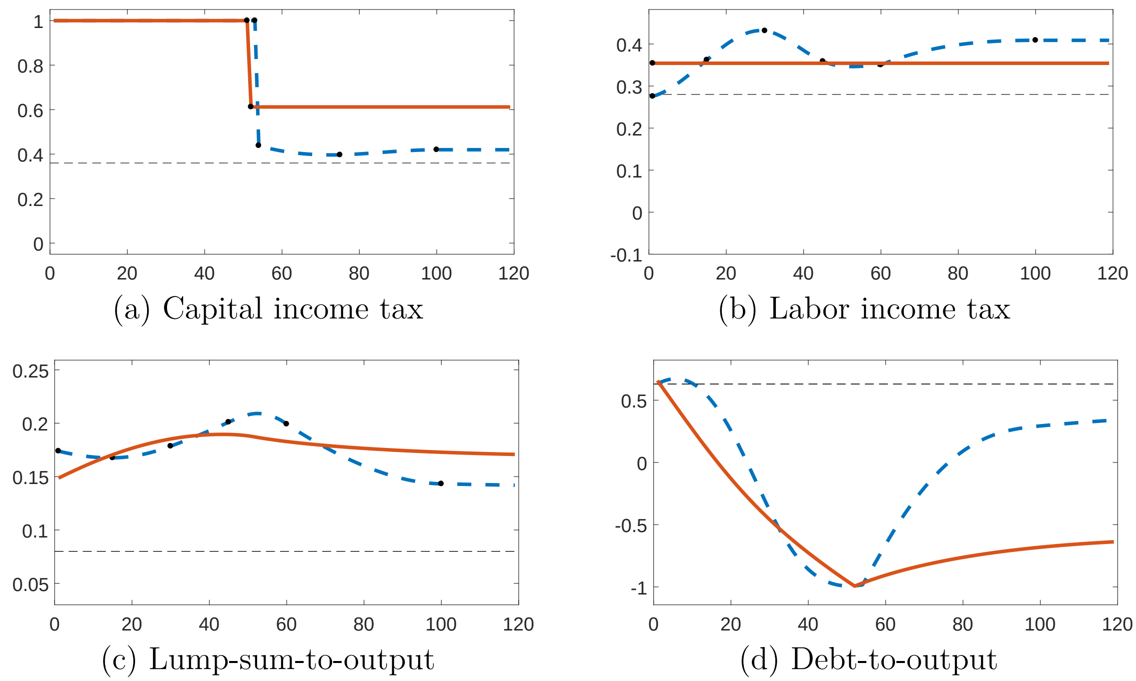


Figure 14: Optimal Fiscal Policy with 4 nodes

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

8.1 Controlling the degree of inequality aversion

The utilitarian welfare function, which we consider in our benchmark results, places equal Pareto weights on every 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

$$W^{\hat{\sigma}} = \left(\int_S \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t v\left(x_t\left(a_0, e^t\right)\right) \right]^{\frac{1-\hat{\sigma}}{1-\sigma}} d\lambda_0 \right)^{\frac{1-\sigma}{1-\hat{\sigma}}},$$

where λ_0 is the initial distribution over individual states (a_0, e_0). Following Benabou (2002), we call $\hat{\sigma}$ the planner's degree of inequality aversion. First note that if $\hat{\sigma} = \sigma$ (the agents' degree of risk aversion), maximizing W^{σ} is equivalent to maximizing the utilitarian welfare function. If $\hat{\sigma} = 0$, then maximizing W^0 is akin to maximizing efficiency as in Section 5.1, that is, the planner has no redistributive concerns and focuses instead in the reduction of distortions and the provision of insurance.[45] Finally, it is easy to see that

$$\lim_{\hat{\sigma}\to\infty} W^{\hat{\sigma}} = \min_{(a_0,e_0)} \mathbb{E}_0 \left[U\left(\left\{x_t\left(a_0,e^t\right)\right\}\right)\right].$$

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 8 displays the results for different levels of $\hat{\sigma}$.

Table 8: Controlling the degree of inequality aversion

	t^*	τ^k	τ^n	T/Y	B/Y	Δ	Δ_L	Δ_I	Δ_R
Degree of Inequality Aversion. Benchmark: $\hat{\sigma}=2$									
$\hat{\sigma}=0.0$	0	53.0	27.5	10.2	40.3	4.6	0.0	0.2	4.5
$\hat{\sigma}=1.0$	35	69.3	31.0	15.6	-45.4	12.8	-2.4	0.9	14.6
$\hat{\sigma}=2.0^*$	51	61.2	35.4	16.8	-59.4	13.2	-4.3	1.8	16.1
$\hat{\sigma}=3.0$	53	61.2	37.5	17.8	-65.4	13.1	-5.0	2.2	16.6
$\hat{\sigma}=4.0$	54	61.5	38.5	18.4	-68.9	13.0	-5.5	2.4	16.7
$\hat{\sigma}=5.0$	55	61.5	39.2	18.6	-70.9	12.9	-5.8	2.6	16.9

Note: When $\hat{\sigma}=2=\sigma$ the welfare function is utilitarian; this is the solution plotted in Figure 14. 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, $t^*=0$. The benchmark result that capital income 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 $\hat{\sigma}\geq 2$, the results do not change significantly with changes in $\hat{\sigma}$. In particular, the final levels of capital and labor income taxes and the composition of the welfare gains are remarkably similar.

8.2 Labor-supply and intertemporal elasticities

One parameter, σ, 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 9 contains the results for other choices of this parameter and also for different levels of Frisch elasticity.

[45] Proposition 8 in Appendix B formalizes this claim.

Table 9: Risk Aversion and Frisch Elasticity (Benchmark: $\sigma = 2$, $\kappa = 0.72$)

	t^*	τ^k	τ^n	T/Y	B/Y	Δ	Δ_L	Δ_I	Δ_R
$\sigma = 1.0$	14	36.0	32.0	11.8	-8.9	3.5	-2.3	0.8	5.0
$\sigma = 3.0$	83	91.8	36.5	21.1	-85.6	22.6	-2.5	1.7	23.5
$\kappa = 0.5$	49	35.2	49.0	22.0	-83.6	23.1	-8.2	4.9	27.7
$\kappa = 1.0$	50	66.5	27.0	13.3	-41.5	11.2	-2.1	0.2	13.5
Benchmark	51	61.2	35.4	16.8	-59.4	13.2	-4.3	1.8	16.1

When σ 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 51 to 14). Moreover, the agents' intertemporal elasticity of substitution increases and their risk aversion is reduced which implies that long-run capital income taxes lead to, at the same time, higher distortions and less benefits. It follows that the optimal long-run capital income tax is lower. This also leads to a higher proportion of welfare gains coming from the level effect and less coming from redistribution. The opposite happens when σ is increased to 3. Intuitively, a higher Frisch elasticity implies a lower optimal labor income tax and a higher associated level effect. Note that these results are in line with the propositions established in Section 1.

8.3 Wealth effects and preference shocks

In the benchmark calibration, productivity shocks affect the amount of labor supplied by the agents, an effect that is magnified by the lack of wealth effects. It is possible to remove this effect and make labor supply independent of the productivity shock, by introducing a concomitant shock to the disutility of labor, i.e. setting $\chi_e = \chi_0 e$. We recalibrate the model under this alternative assumption and compute the optimal policy which we present in Table 10. Note that the optimal policy and welfare decomposition are very similar. The main difference being the lower long-run capital income taxes and higher t^*. As discussed in Section 4.4, variations in these features of capital income taxes are close substitutes.

Table 10: Benchmark versus Calibration with Preference Shocks

	t^*	τ^k	τ^n	T/Y	B/Y	Δ	Δ_L	Δ_I	Δ_R
Pref. Shocks	57	29.6	40.4	16.5	-74.4	16.0	-7.2	2.5	21.9
Benchmark	51	61.2	35.4	16.8	-59.4	13.2	-4.3	1.8	16.1

9 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 and labor income at levels roughly consistent with the prevailing ones in the US. Moreover, we show that to achieve redistribution, it is optimal for the planner to set capital income taxes at the imposed upper bound of 100 percent for several years. By decomposing the welfare gains we diagnose that, relative to the current US tax system, this policy exacerbates the amount of distortions to agents' decisions. On the other hand, it leads to a substantial amount of redistribution and insurance, with the former being significantly more relevant for the welfare gains associated with the optimal policy.

Finally, we do not view our results as a final answer to our initial question: "how should governments conduct fiscal policy in the presence of inequality and individual risk?" 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. The model also abstracts from international financial markets; capital income taxes as high as the ones we find optimal in this paper are unlikely to survive if agents are able to move their assets overseas. We also abstract from life-cycle issues and have a relatively simple tax structure. Our method, however, could be used to approximate the solution to Ramsey problems in more elaborate models, the main constraint being computational power and how long it would take to compute one transition.

References

Acikgoz, Omer, Marcus Hagedorn, Hans Holter, and Yikai Wang, "The Optimum Quantity of Capital and Debt," Technical Report May 2018.

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 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.

Bakis, Ozan, Baris Kaymak, and Markus Poschke, "Transitional Dynamics and the Optimal Progressivity of Income Redistribution," *Review of Economic Dynamics*, July 2015, *18* (3), 679–693.

Barsky, Robert B, N Gregory Mankiw, and Stephen P Zeldes, "Ricardian Consumers with Keynesian Propensities," *American Economic Review*, September 1986, *76* (4), 676–691.

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 J. Sargent, "Public debt in economies with heterogeneous agents," *Journal of Monetary Economics*, 2017, *91* (C), 39–51.

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.

Cesarini, David, Erik Lindqvist, Matthew J. Notowidigdo, and Robert Östling, "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries," *American Economic Review*, December 2017, *107* (12), 3917–46.

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.

Chari, V. V., Juan Pablo Nicolini, and Pedro Teles, "Optimal Capital Taxation Revisited," Working Papers 752, Federal Reserve Bank of Minneapolis July 2018.

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

Chen, Yunmin, YiLi Chien, and C.C. Yang, "Implementing the Modified Golden Rule? Optimal Ramsey Capital Taxation with Incomplete Markets Revisited," Working Papers, Federal Reserve Bank of St. Louis January 2018.

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.

Chien, YiLi and Yi Wen, "Optimal Ramsey Capital Income Taxation —A Reappraisal," Working Papers 2017-24, Federal Reserve Bank of St. Louis August 2017.

Civale, Simone, Luis Diez-Catalan, and Fatih Fazilet, "Discretizing a Process with Non-Zero Skewness and High Kurtosis," Technical Report April 2016.

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.

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

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

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.

Greulich, Katharina, Sarolta Laczó, and Albert Marcet, "Pareto-Improving Optimal Capital and Labor Taxes," Working Papers 887, Barcelona Graduate School of Economics March 2016.

Guvenen, Fatih, "Macroeconomics With Heterogeneity: A Practical Guide," Working Paper 17622, National Bureau of Economic Research November 2011.

—, Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?," NBER Working Papers 20913, National Bureau of Economic Research, Inc January 2015.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, May 2009, *1* (1), 319–354.

—, —, and —, "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," *The Quarterly Journal of Economics*, 2017, *132* (4), 1693–1754.

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.

Itskhoki, Oleg and Benjamin Moll, "Optimal Development Policies with Financial Frictions," Technical Report July 2018.

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.

Kan, Alexander Rinnooy and G. T. Timmer, "Stochastic Global Optimization Methods, Part I, Clustering Methods," *Mathematic Programming*, 1987, *39*, 27–56.

— and —, "Stochastic Global Optimization Methods, Part II, Multilevel Methods," *Mathematic Programming*, 1987, *39*, 57–78.

Krueger, Dirk and Alexander Ludwig, "On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium," *Journal of Monetary Economics*, 2016, *77*, 72 – 98. "Inequality, Institutions, and Redistribution"held at the Stern School of Business,New York University,April 24-25, 2015.

— **and** — , "Optimal Taxes on Capital in the OLG Model with Uninsurable Idiosyncratic Income Risk," CEPR Discussion Papers 12717, C.E.P.R. Discussion Papers February 2018.

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.

Nardi, Mariacristina De and Giulio Fella, "Saving and Wealth Inequality," *Review of Economic Dynamics*, October 2017, *26*, 280–300.

Nuño, Galo and Carlos Thomas, "Optimal monetary policy with heterogeneous agents," Working Papers 1624, Banco de España;Working Papers Homepage October 2016.

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

Powell, Michael, "The BOBYQA algorithm for bound constrained optimization without derivatives," Technical Report, Department of Applied Mathematics and Theoretical Physics, Cambridge University June 2009.

Ragot, Xavier and Francois Le Grand, "Optimal Fiscal Policy with Heterogeneous Agents and Aggregate Shocks," Technical Report 2017.

Röhrs, Sigrid and Christoph Winter, "Reducing government debt in the presence of inequality," *Journal of Economic Dynamics and Control*, 2017, *82* (C), 1–20.

Rouwenhorst, Geert, "Asset Pricing Implications of Equilibrium Business Cycle Models," 1995. in Frontiers of Business Cycle Research, edited by Thomas F. Cooley. Princeton University Press.

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.

Tauchen, George, "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 1986, *20* (2), 177 – 181.

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

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

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 on our websites.[46]

A Proofs for two-period economies

A.1 Risk economy

Define $\tau_R^k \equiv r\tau^k / (1+r)$. Six equations determine a tax distorted equilibrium $(K, n_L, n_H, r, w;$ τ^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

$$r = f_K(K,N), \quad \text{and} \quad w = f_N(K,N), \quad \text{where } N = \pi e_L n_L + (1-\pi) e_H n_H \tag{9.1}$$

and the government's budget constraint. Using equation (9.1) 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

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

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[47]):

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

where Θ^n and Θ^k are complicated functions of equilibrium variables.[48]

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.

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

[47] Mathematica codes that compute all the algebraic steps are available on our websites.

[48] Here are the exact formulas:

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

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

Proof of Proposition 1. First note 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. $\Theta^k = 0$ 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 τ_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 τ^n we obtain equation (1.1) and substituting it back in the equation for τ_R^k we obtain $\tau_R^k = 0$; and, therefore, $\tau^k = 0$. This is the only pair $\left(\tau^n, \tau_R^k\right) \in (-\infty, 1)^2$ 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

Proof of Proposition 3. First note that without risk $e^t = \{e_0\}$, so that

$$\lambda_t \left(a_0, e^t\right) = \lambda_0 \left(a_0, e_0\right), \quad \text{and} \quad x_t \left(a_0, e^t\right) = x_t \left(a_0, e_0\right),$$

for all (a_0, e_0). It follows from (4.4) that,

$$U\left(\left\{\bar{x}_t \left(a_0, e_0\right)\right\}\right) = U\left(\left\{x_t \left(a_0, e_0\right)\right\}\right),$$

and, therefore,

$$\eta \left(a_0, e_0\right) = 1,$$

where

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

for all (a_0, e_0). Hence, we obtain

$$
\begin{aligned}
U\left(\left\{\bar{X}_t\right\}\right) &= \sum_{t=0}^{\infty} \beta^t u\left(\bar{X}_t\right) \overset{(4.6)}{=} \sum_{t=0}^{\infty} \beta^t v\left(\int \bar{x}_t\left(a_0, e_0\right) d\lambda_t\left(a_0, e_0\right)\right) \\
&\overset{(4.5)}{=} \sum_{t=0}^{\infty} \beta^t v\left(\int x_t\left(a_0, e^t\right) d\lambda_t\left(a_0, e^t\right)\right) = \sum_{t=0}^{\infty} \beta^t v\left(X_t\right) = U\left(\left\{X_t\right\}\right),
\end{aligned}
$$

which establishes the result. Next, without inequality, we have that $\bar{x}_t^j\left(a_0, e_0\right) = \bar{X}_t$ for all t and all (a_0, e_0). ■

Proof of Proposition 4. Note that

$$
U\left(a\left\{x_t\right\}\right) = a^{1-\sigma} U\left(\left\{x_t\right\}\right). \tag{9.2}
$$

Suppressing the dependence on (a_0, e_0), it follows that

$$
\begin{aligned}
\int \mathbb{E}_0\left[U\left(\left\{x_t^R\right\}\right)\right] d\lambda_0 &\overset{(4.4)}{=} \int U\left(\left\{\bar{x}_t^R\right\}\right) d\lambda_0 \overset{(4.8)}{=} U\left(\left(1-p_{ineq}^R\right)\left\{\bar{X}_t^R\right\}\right) \overset{(9.2)}{=} \left(1-p_{ineq}^R\right)^{1-\sigma} U\left(\left\{\bar{X}_t^R\right\}\right) \\
&\overset{(4.7)}{=} \left(1-p_{ineq}^R\right)^{1-\sigma} U\left(\left(1-p_{risk}^R\right)\left\{X_t^R\right\}\right) \overset{(9.2)}{=} \left(\left(1-p_{ineq}^R\right)\left(1-p_{risk}^R\right)\right)^{1-\sigma} U\left(\left\{X_t^R\right\}\right) \\
&\overset{(4.3)}{=} \left(\left(1-p_{ineq}^R\right)\left(1-p_{risk}^R\right)\right)^{1-\sigma} U\left(\left(1+\Delta_L\right)\left\{X_t^{NR}\right\}\right) \\
&\overset{(9.2)}{=} \left(\left(1+\Delta_L\right)\left(1-p_{ineq}^R\right)\left(1-p_{risk}^R\right)\right)^{1-\sigma} U\left(\left\{X_t^{NR}\right\}\right) \\
&\overset{(9.2)}{=} \left(\left(1+\Delta_L\right)\left(1-p_{ineq}^R\right)\frac{\left(1-p_{risk}^R\right)}{\left(1-p_{risk}^{NR}\right)}\right)^{1-\sigma} U\left(\left(1-p_{risk}^{NR}\right)\left\{X_t^{NR}\right\}\right) \\
&\overset{(4.7)}{=} \left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\left(1-p_{ineq}^R\right)\right)^{1-\sigma} U\left(\left\{\bar{X}_t^{NR}\right\}\right) \\
&\overset{(9.2)}{=} \left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\frac{\left(1-p_{ineq}^R\right)}{\left(1-p_{ineq}^{NR}\right)}\right)^{1-\sigma} U\left(\left(1-p_{ineq}^{NR}\right)\left\{\bar{X}_t^{NR}\right\}\right) \\
&\overset{(4.8)}{=} \left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\left(1+\Delta_R\right)\right)^{1-\sigma} \int U\left(\left\{\bar{x}_t^{NR}\right\}\right) d\lambda_0 \\
&\overset{(4.7)}{=} \left(\left(1+\Delta_L\right)\left(1+\Delta_I\right)\left(1+\Delta_R\right)\right)^{1-\sigma} \int \mathbb{E}_0\left[U\left(\left\{x_t^{NR}\right\}\right)\right] d\lambda_0 \\
&\overset{(9.2)}{=} \int \mathbb{E}_0\left[U\left(\left(1+\Delta_R\right)\left(1+\Delta_I\right)\left(1+\Delta_L\right)\left\{x_t^{NR}\right\}\right)\right] d\lambda_0.
\end{aligned}
$$

The result, then, follows from the definition of Δ in equation (4.2). ■

Proposition 8 *If the certainty equivalents are constant over time, i.e.* $\bar{x}_t^j\left(a_0, e_0\right) = \bar{x}^j\left(a_0, e_0\right)$ *for* $j = R, NR$, *then, maximizing* $W^0 = \left(\int \mathbb{E}_0\left[U\left(\left\{x_t\left(a_0, e^t\right)\right\}\right)\right]^{\frac{1}{1-\sigma}} d\lambda_0\left(a_0, e_0\right)\right)^{1-\sigma}$ *is equivalent to maximizing* $\left(1+\Delta_L\right)\left(1+\Delta_I\right)$.

Proof. First note that, for $j = R, NR$,

$$\mathbb{E}_0 \left[U \left(\left\{ x_t^j \right\} \right) \right]^{\frac{1}{1-\sigma}} \overset{(4.4)}{=} U \left(\left\{ \bar{x}^j \right\} \right)^{\frac{1}{1-\sigma}} = \left(\sum_{t=0}^{\infty} \beta^t \left(\bar{x}^j \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \left(\frac{1}{1-\beta} \right)^{\frac{1}{1-\sigma}} \bar{x}^j, \tag{9.3}$$

and, therefore

$$\begin{aligned}
\int \mathbb{E}_0 \left[U \left(\left\{ x_t^R \right\} \right) \right]^{\frac{1}{1-\sigma}} d\lambda_0 \overset{(9.3)}{=} \int \left(\frac{1}{1-\beta} \right)^{\frac{1}{1-\sigma}} \bar{x}^R d\lambda_0 &\overset{(4.6)}{=} \left(\frac{1}{1-\beta} \right)^{\frac{1}{1-\sigma}} \bar{X}^R = U \left(\left\{ \bar{X}^R \right\} \right)^{\frac{1}{1-\sigma}} \\
&\overset{(4.7)}{=} U \left(\left(1 - p_{risk}^R \right) \left\{ X_t^R \right\} \right)^{\frac{1}{1-\sigma}} = \left(1 - p_{risk}^R \right) U \left(\left\{ X_t^R \right\} \right)^{\frac{1}{1-\sigma}} \\
&\overset{(4.3)}{=} \left(1 - p_{risk}^R \right) U \left(\left(1 + \Delta_L \right) \left\{ X_t^{NR} \right\} \right)^{\frac{1}{1-\sigma}} \\
&= \frac{\left(1 - p_{risk}^R \right)}{\left(1 - p_{risk}^{NR} \right)} \left(1 + \Delta_L \right) U \left(\left(1 - p_{risk}^{NR} \right) \left\{ X_t^{NR} \right\} \right)^{\frac{1}{1-\sigma}} \\
&\overset{(4.7)}{=} \left(1 + \Delta_I \right) \left(1 + \Delta_L \right) U \left(\left\{ \bar{X}^{NR} \right\} \right)^{\frac{1}{1-\sigma}} \\
&= \left(1 + \Delta_I \right) \left(1 + \Delta_L \right) \left(\frac{1}{1-\beta} \right)^{\frac{1}{1-\sigma}} \bar{X}^{NR} \\
&\overset{(4.6)}{=} \left(1 + \Delta_I \right) \left(1 + \Delta_L \right) \int \left(\frac{1}{1-\beta} \right)^{\frac{1}{1-\sigma}} \bar{x}^{NR} d\lambda_0 \\
&\overset{(9.3)}{=} \left(1 + \Delta_I \right) \left(1 + \Delta_L \right) \int \mathbb{E}_0 \left[U \left(\left\{ x_t^{NR} \right\} \right) \right]^{\frac{1}{1-\sigma}} d\lambda_0
\end{aligned}$$

which establishes the result. ■

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, where we present more detailed versions of the proofs. 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 7. For simplicity, we assume that agents are divided into a finite number of types $i \in I$ of relative size π_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 τ^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

$$\sum_{t=0}^{\infty} p_t \left((1+\tau^c) \, c_{i,t} - (1-\tau_t^n) \, w_t e_i n_{i,t} \right) \leq R_0 a_{i,0} + T. \tag{9.4}$$

Similarly, the government's budget constraint simplifies to

$$R_0 B_0 + T + \sum_t p_t G = \sum_t p_t \left(\tau^c C_t + \tau_t^n w_t N_t + \tau_t^k r_t K_t \right). \tag{9.5}$$

The resource constraint is given by

$$C_t + G + K_{t+1} = f\left(K_t, N_t\right), \quad \text{for all } t \geq 0. \tag{9.6}$$

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 (9.4) 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 (9.5) holds; and (iv) markets clear: the resource constraints (9.6) hold.*

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

$$U\left(C_t, N_t; \varphi\right) \equiv \max_{c_{i,t}, n_{i,t}} \sum_i \pi_i \varphi_i u\left(c_{i,t}, n_{i,t}\right) \quad \text{s.t.} \ \sum_i \pi_i c_{i,t} = C_t \quad \text{and} \quad \sum_i \pi_i e_i n_{i,t} = N_t \tag{9.7}$$

where u is given by equation (3.1), for some vector $\varphi \equiv \{\varphi_i\}$ of market weights $\varphi_i \geq 0$. Let $c_{i,t}^m\left(C_t, N_t; \varphi\right)$, and $n_{i,t}^m\left(C_t, N_t; \varphi\right)$ be the argmax of this problem. It can be shown that[49]

[49]Where constants are defined as follows:

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

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

Then, implementability constraints can be written as

$$
\begin{aligned}
&\sum_{t=0}^{\infty} \beta^t \left(U_C\left(C_t, N_t; \varphi\right) c_{i,t}^{m}\left(C_t, N_t; \varphi\right) + U_N\left(C_t, N_t; \varphi\right) e_i n_{i,t}^{m}\left(C_t, N_t; \varphi\right)\right) \\
&\qquad = U_C\left(C_0, N_0; \varphi\right)\left(\frac{R_0 a_{i,0} + T}{1+\tau^c}\right) \quad \text{for all } i \in I
\end{aligned} \tag{9.8}
$$

Proposition 9 *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 (9.6) hold and there exist market weights φ and a lump-sum tax T so that the implementability conditions (9.8) hold for all $i \in I$. Individual allocations can then be computed using functions $c_{i,t}^m$ and $n_{i,t}^m$, and 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,

$$
\sum_{t=0}^{\infty} \beta^t \pi_i \lambda_i u\left(c_{i,t}, n_{i,t}\right), \tag{9.9}
$$

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 $t \geq 1$.

Note 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

$$
U_C(t) \geq \beta U_C(t+1), \tag{9.10}
$$

is equivalent to $\tau_{t+1}^k \leq 1$. Moreover, note that τ_0^k and T_0 have not been substituted out of the implementability constraint. The fact that τ_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

$$
N_0 = \bar{N}_0, \tag{9.11}
$$

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

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

Finally, we can use Proposition 9 to rewrite the Ramsey problem as choosing $\{C_t, N_{t+1}, K_{t+1}\}_{t=0}^{\infty}$, T, and φ to maximize (9.9) subject to (9.6) for all $t \geq 0$, (9.8) for all $i \in I$ with multiplier μ_i, (9.10) for all $t \geq 0$ with multiplier η_t, and (9.11). 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 (9.6) for all $t \geq 0$, (9.10) for all $t \geq 0$, and (9.11), 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[50]

$$\begin{aligned} W\left(C_t, N_t; \varphi, \mu, \lambda\right) = & \frac{1}{1-\sigma}\left(C_t - \Omega^n \chi \frac{\kappa}{1+\kappa}\left(N_t\right)^{\frac{1+\kappa}{\kappa}}\right)^{-\sigma} \\ & * \left(\Phi C_t - \left(\Phi + (1-\sigma)\Psi\right)\Omega^n \chi \frac{\kappa}{1+\kappa}\left(N_t\right)^{\frac{1+\kappa}{\kappa}}\right) \end{aligned} \tag{9.12}$$

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)

[50] Where constants are defined as follows:

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

together with equilibrium conditions imply the following equations[51]

$$\sum_i \pi_i \mu_i = 0 \tag{9.13}$$

$$\frac{\tau_t^n + \tau^c}{1+\tau^c} = \frac{\Psi \Theta_t}{\Phi \Theta_t + \Psi\left(\sigma + \Theta_t\right) + \Upsilon_t \sigma\left(\beta \eta_{t-1} - \eta_t\right)}, \text{ for } t \geq 1 \tag{9.14}$$

$$\frac{R_{t+1}}{R_{t+1}^*} = \frac{\Phi \Theta_{t+1} + \Psi \sigma + \Upsilon_{t+1} \sigma\left(\beta \eta_t - \eta_{t+1}\right)}{\Phi \Theta_t + \Psi \sigma + \Upsilon_t \sigma\left(\beta \eta_{t-1} - \eta_t\right)} \frac{\Theta_t}{\Theta_{t+1}}, \text{ for } t \geq 0 \tag{9.15}$$

Note 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 Ψ that

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

where the last equality follows from equation (9.13). Since $\Psi = 0$, it follows from equation (9.12) that

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

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 5. Using Lemma 3, from equation (9.14) 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 (9.15) that $\tau_1^k < 1$ if

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

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

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

from (9.15) that $\tau^k_{t^*} < 1$ 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 $\tau^k_t = 0$ for $t \geq t^*+1$. The result follows from the fact that η_{-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 6. 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

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

where the last equality follows from equation (9.13). Since here we assume that τ^k_t does not have to be bounded by 1, it follows that $\eta_t = 0$ for all $t \geq 1$. Then, equation (7.1) follows directly from equation (9.14), (7.2) from its derivative with respect to Θ_t, and (7.3) from equations (9.14) and (9.15). ■

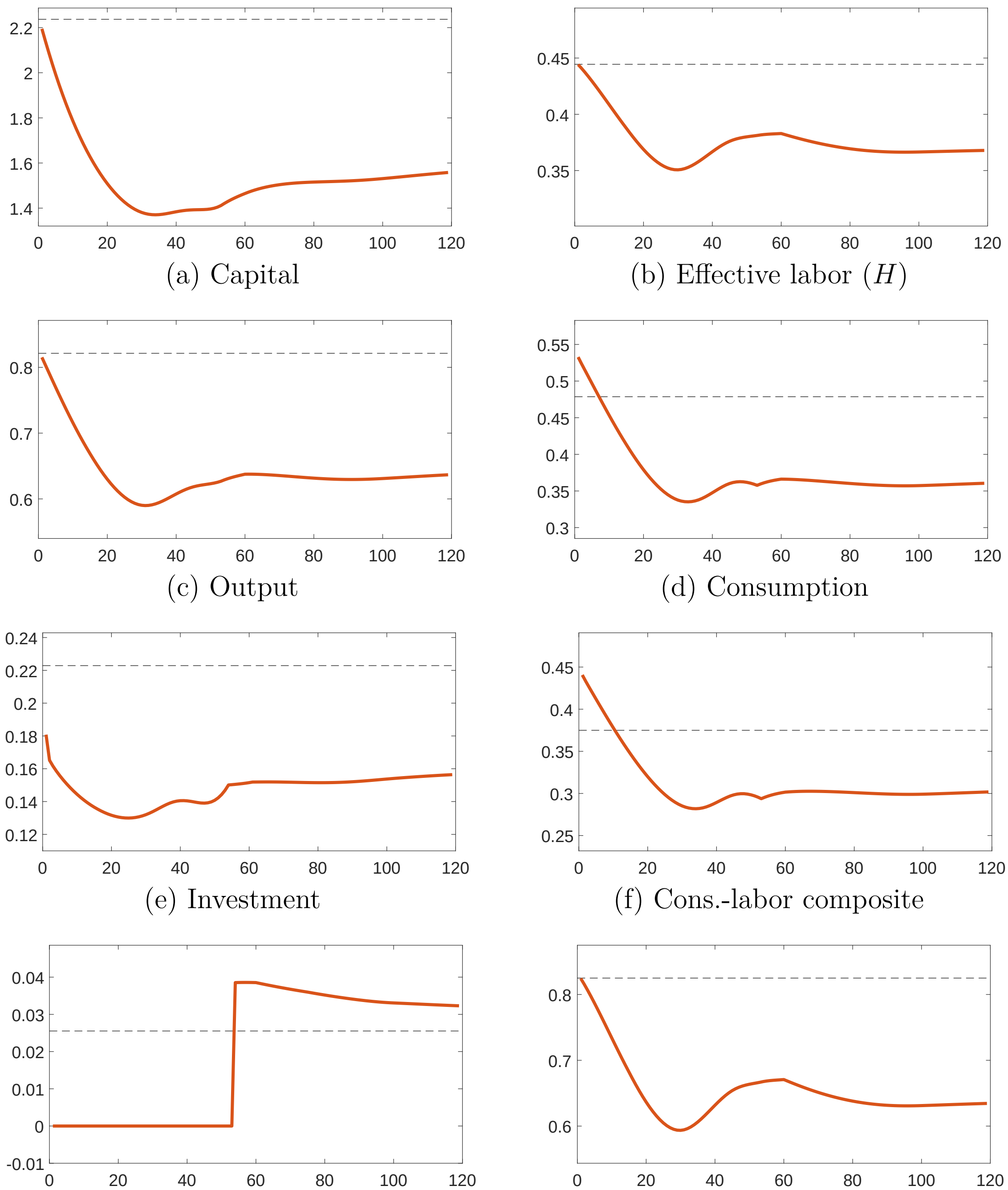
C.3 Economy 4

Proof of Proposition 7. Equation (7.3) can be established for all $t \geq 1$, by substituting (9.14) into (9.15). 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 5, which we suppress here.[52] Hence, for $t \geq t^*$ we can obtain τ^n_t by using (7.1), which follows from (9.14) with $\eta_t = 1$. For the same time period τ^k_t can then be found by using (7.3). Now, having $\tau^n_{t^*}$ we can use the fact that $\tau^k_t = 1$ and (7.3) moving backwards to obtain τ^n_t for $t < t^*$. ■

[52] 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 composite

(g) After-tax int. rates (h) After-tax wages

Figure 15: Aggregates: Benchmark

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