Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks

Dyrda, S.; Pedroni, M.

DOI
10.2139/ssrn.3289306

Publication date
2020

Document Version
Submitted manuscript

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: December 16, 2020

Abstract

We evaluate the optimal fiscal policy in a standard incomplete-markets model with uninsurable idiosyncratic shocks, where a Ramsey planner chooses time-varying paths of proportional capital and labor income taxes, lump-sum transfers (or taxes), and government debt. We find that: (1) short-run capital income taxes are effective in providing redistribution since the tax base is relatively unequal and inelastic; (2) an increasing pattern of labor income taxes over time mitigates intertemporal distortions from capital income taxes; (3) the optimal policy expands the US social welfare system significantly, increasing overall transfers by roughly 50 percent; (4) two thirds of the welfare gains come from redistribution and the remaining third come mostly from insurance; and (5) redistribution also leads to a more efficient allocation of labor via wealth effects on labor supply—lower productivity households can afford to work relatively less.

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. Acknowledgments: We thank Anmol Bhandari, V.V. Chari, Carlos Eugênio da Costa, Jonathan Heathcote, Hans Holter, Albert Jan Hummel, Larry Jones, Gueorgui Kambourov, Burhan Kuruscu, Chris Phelan, Diego Restuccia, José-Víctor Ríos-Rull, Cezar Santos, Christian Stoltenberg, 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. Computations were also performed on the Niagara supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation; the Government of Ontario; Ontario Research Fund–Research Excellence; and the University of Toronto.

1 Introduction

How and to what extent should fiscal policy be used to mitigate household inequality and risk? We provide a quantitative answer to these questions by studying a Ramsey problem in the standard incomplete-markets (SIM) model, a general equilibrium model with heterogeneous agents and uninsurable idiosyncratic labor income risk.[1]

We begin with a detailed calibration of the SIM model that replicates several aspects of the US economy, including the cross-sectional distribution of wealth, earnings, hours worked, consumption, and total income, as well as statistical properties of the labor income process of households. We then consider a Ramsey planner that finances an exogenous stream of government expenditures with proportional capital and labor income taxes, lump-sum transfers (or taxes), and government debt. We allow policy to be *time varying* and evaluate the welfare function over the *transition* path. To solve for the optimal paths of fiscal instruments, we parameterize them in the time domain using flexible polynomials, then maximize welfare using a global optimization algorithm.

For our benchmark calibration, we find that a utilitarian planner would tax capital income heavily in the short run, but also at a positive rate of 27 percent in the long run. These long-run taxes are lower than the prevailing ones in the US, which we calculate to be about 42 percent. Labor income taxes increase steeply in the short-run, reaching 39 percent in the long run, higher than the prevailing rate of 23 percent. These changes in taxes are then used to finance an increase in lump-sum transfers of roughly 50 percent on average over time. The ratio of government debt to GDP more than doubles to 154 percent in the long run. This policy leads to welfare gains equivalent to a permanent increase in consumption of 3.5 percent.

More generally, we provide new insights about the dynamics of the optimal policy in the SIM model. High short-run capital income taxes are effective in providing redistribution, since the tax base is relatively unequal and inelastic. Labor income taxes are increasing over this period to mitigate the corresponding intertemporal distortions. Lump-sum transfers are front-loaded allowing households to move away from their borrowing constraints. In the long-run, positive capital and labor income taxes provide insurance for households' long-run risk. These qualitative features of the optimal policy are robust to significant changes to the calibration of the model.

To disentangle the main forces that determine the optimal policy, we develop a procedure to decompose welfare gains into what comes from the reduction of distortions to households'

[1] Originally developed by Bewley (1986), Imrohoruglu (1989), Huggett (1993), and Aiyagari (1994).

decisions, from redistribution (the reduction of ex-ante risk) and from insurance (the reduction of ex-post risk). The welfare gains of 3.5 percent from the optimal policy can be decomposed into: (i) 0.2 percent from a reduction in distortions, (ii) 1.2 percent from insurance, and (iii) 2.1 percent from redistribution. This decomposition is particularly useful when considering policy variations since it allows us to measure the effects on each of these components separately.

These components of welfare must be considered on balance in the design of the optimal policy. Capital and labor income are both unequally distributed between households and risky over time. Labor and capital income taxes distort households' savings and labor supply decisions, but rebating their revenue via lump-sum transfers effectively provides redistribution and insurance. We formalize and quantify this trade-off by: (1) analytically characterizing the optimal policy in a two-period version of the SIM model; (2) considering perturbations to the optimal policy and quantifying their implications for distortions, inequality, and risk; and (3) measuring the effect of varying the intertemporal elasticity of substitution and Frisch elasticity on optimal taxes.

To investigate further the determinants of the optimal policy, we also consider a Ramsey planner that disregards equality concerns and focuses only on efficiency (i.e. minimizing distortions and risk). The optimal policy in this case is remarkably similar to the benchmark utilitarian one. This is particularly surprising since redistribution accounts for the largest share of the welfare gains in the benchmark results. The reason for this is that redistribution is actually complementary to efficiency. Transferring resources from rich/productive households to poor/unproductive ones leads, through wealth effects on labor supply, to a relative increase in hours worked by the more productive. The end result is a substantial increase in average labor productivity. This effect is strong enough that it is optimal to provide a considerable amount of redistribution even if the sole purpose is to maximize efficiency.

We also show that disregarding transitional welfare effects or the time variation of fiscal instruments can be misleading. To make this point, we first compute the stationary fiscal policy that maximizes steady-state welfare. We show it is very different from the optimal. Further, implementing it from the beginning and accounting for transitional effects, the policy would actually lead to a welfare *loss* equivalent to a 3.5 percent permanent reduction in consumption. This is a result of abstracting from the costs associated with accumulating the eventual long-run level of capital and, more importantly, the distributional effects of imposing the policy on the current population. Further, if transition is taken into account, but fiscal instruments are restricted to being constant over time, the welfare gains are roughly half of the optimal. Allowing front-loading of capital income taxes then generates 80 percent of the optimal gains, so a relatively simple time variation already delivers the lion's share of the welfare gains. However, allowing further movements in the paths of fiscal instruments, besides delivering the extra 20 percent of welfare gains, is important for determining long-run optimal tax levels and other properties of the long-run Ramsey allocation.

To illustrate the role of market incompleteness and highlight why and how our results differ from the existing complete-markets Ramsey literature, we consider complete-markets versions of our model in which we can analytically characterize the optimal fiscal policy. In a representative-agent economy without any heterogeneity, it is optimal to obtain all necessary revenue via lump-sum taxes. Heterogeneity in labor productivity rationalizes distortive labor income taxes for redistributive purposes. Similarly, asset heterogeneity leads to high initial capital income taxes that go to zero after a finite number of periods; in the short run with high capital income taxes, labor income taxes are increasing over time to mitigate intertemporal distortions. If both types of heterogeneity are present, the over-time pattern of optimal capital and labor income taxes is qualitatively and quantitatively similar to those from the SIM model with the notable exception that long-run capital income taxes are positive in the SIM model. Hence, long-run capital income taxes in the SIM model are used to provide insurance for the privately uninsurable risk that is present when markets are incomplete.

In the complete-markets model, the timing of lump-sum transfers and the corresponding path of government debt is indeterminate since the Ricardian equivalence holds. In the SIM model, this is not the case. We find that it is optimal to mostly front-load lump-sum transfers so that households move away from their borrowing constraints, which allows them to better absorb income shocks. With few constrained households in the long run, the effect of changes to long-run debt on household decisions is minor. This implies that long-run debt levels have secondary welfare implications, especially for a planner that discounts the future.

Related Literature

Aiyagari (1995) provides a rationale for positive long-run capital income taxes in the SIM model: these taxes implement the modified golden rule by attenuating households precautionary savings.[2] We quantify, in particular, the specific value for the optimal long-run capital income taxes. Acikgoz (2015) and, more recently, Acikgoz, Hagedorn, Holter, and Wang (2018) obtained additional long-run optimality conditions.[3] We extend the results

[2] Chamley (2001) provides a complementary rationale, transferring from the rich to the poor in the long-run is Pareto improving since, far enough in the future, everyone has the same probability of being in either condition. Chen, Yang, and Chien (2020) argue that the existence of the Ramsey steady state, assumed by Aiyagari (1995), depends on the value of intertemporal elasticity of substitution.

[3] Acikgoz et al. (2018) also argue that in the SIM model, long-run fiscal policy can be characterized inde-

from Acikgoz et al. (2018) to obtain long-run optimality conditions for the balanced-growth-path preferences we use and show that our results do satisfy these conditions. We find this to be reassuring about the accuracy of both methods.

Gottardi, Kajii, and Nakajima (2015) and Heathcote, Storesletten, and Violante (2017) analytically characterize the optimal fiscal policy in stylized versions of the SIM model. Krueger and Ludwig (2018) do the same in an overlapping generations setup. Their approaches lead to elegant and insightful closed-form solutions. We take a more quantitative approach which allows us to match some aspects of the data, in particular measures of inequality and risk, which we find to be important for the determination of the optimal tax system.

There is a limited but growing literature on Ramsey problems in quantitative frameworks with heterogeneity. Itskhoki and Moll (2019) study optimal dynamic development policies in an incomplete-markets model where heterogeneous producers are subject to financial frictions. Nuño and Thomas (2016) use a novel continuous-time technique to solve for optimal monetary policy, including optimal transition, in a version of the SIM model with money. Ragot and Grand (2020) solve the Ramsey problem in the SIM model with aggregate technology shocks by truncating the histories of idiosyncratic shocks. Our contribution to this literature is to develop a technique for solving Ramsey problems which can be applied to a wide range of models including a realistically calibrated SIM model. Also, our welfare decomposition offers a clean way of breaking down welfare gains in non-stationary environments with heterogeneity and risk.

There is a larger literature analyzing optimal policy in the steady state—for instance, Conesa, Kitao, and Krueger (2009)—or optimal constant policy including transitional effects—Bakis, Kaymak, and Poschke (2015),Krueger and Ludwig (2016) and Boar and Midrigan (2020). To our knowledge, Domeij and Heathcote (2004) were the first to quantify the importance of accounting for transitional effects of fiscal policy in the SIM model, showing that the short-run distributional losses that result from reducing capital income taxes dominate the long-run gains.[4] We show that, in our framework, it is important to not only account for transitional effects but also to allow policy instruments to change over time.

We also contribute to the literature on the interaction between government-debt policy and market incompleteness. In an influential paper, Aiyagari and McGrattan (1998) show that

pendently of initial conditions and solve backwards for the optimal transition. We discuss in detail the relationship between our method and results and theirs in Appendix I.

[4]Huggett (1997) developed an algorithm to compute transition in the SIM model, and Conesa and Krueger (1999) account for transitional effects of social-security policies in an overlapping-generations version o the SIM model.

current levels of debt-to-output are close to the level that maximizes steady-state welfare. Röhrs and Winter (2017) show that calibrating the model to match inequality measures leads to high levels of government assets being optimal.[5] We target cross-sectional statistics and properties of the labor income process, and compute optimal government debt not only in the long run but also in transition. We then quantify the importance of time-varying debt under optimal policy in the SIM model.

Finally, there is an extensive literature on Ramsey problems in complete-markets economies. The most well-known result, due to Judd (1985) and Chamley (1986), that capital income taxes should converge to zero in the long run[6] has been refined by Straub and Werning (2020). It remains true in the complete-markets version of our model since we allow for lump-sum taxes.[7] Werning (2007) characterizes optimal policy for this class of economies allowing for complete expropriation of initial capital holdings. We extend that characterization to impose an upper bound on capital income taxes and obtain complete-markets results that are comparable to our benchmark results. Following a numerical approach similar to ours, Conesa and Garriga (2008) use flexible time-dependent instruments to study social security reform. Bassetto (2014), Saez and Stantcheva (2018), and Greulich, Laczó, and Marcet (2019) also study optimal fiscal policy with heterogeneous households focusing on different dimensions.

2 Mechanism: Two-Period Economy

In this section, we characterize the optimal tax system in a two-period version of the quantitative model used below. In it, the government is allowed to finance all necessary revenue with non-distortive lump-sum taxes, but chooses instead to use distortive capital and labor income taxes. We explore the effects of risk and inequality in turn.

[5] Bhandari, Evans, Golosov, and Sargent (2017) investigate the role of government debt in an incomplete markets economy with fixed heterogeneity and aggregate risk. They highlight that having some households borrowing constrained can be beneficial since it magnifies price effects of changes in government debt. This mechanism plays a role in some of our results.

[6] Among others, Jones, Manuelli, and Rossi (1997), Atkeson, Chari, and Kehoe (1999) and Chari, Nicolini, and Teles (2018) show this result is robust to a relaxation of a number of assumptions.

[7] We discuss this in detail in Appendix F.8.

2.1 The effect of risk

Consider an economy with a measure one of ex-ante identical households who live for two periods. Suppose the period utility function is given by

$$u(c,h) = \frac{(c^{\gamma}(1-h)^{1-\gamma})^{1-\sigma}}{1-\sigma}, \tag{2.1}$$

where c and h are the levels of consumption and labor, γ controls the consumption share, and σ controls the preference for risk and over-time smoothness. Also, suppose that households discount the future by a factor of β.

In period 1, each household receives an endowment of ω consumption goods, which can be invested into a risk-free asset a, and supplies $\bar{h}$ units of labor inelastically. In period 2, households receive income from the asset they saved in period 1 and from labor. Labor is supplied endogenously in period 2. The productivity of the labor supplied is random and can take two values: e_L with probability π_L, and $e_H > e_L$ with probability π_H, with the mean productivity normalized to 1. These productivity shocks are independent across consumers, and a law of large numbers operates so that the fraction of households with each productivity level equals their probability.

In period 2, output is produced using capital, K, and labor, N, and a constant-returns-to-scale neoclassical production function $F(K,N)$ which includes undepreciated capital. The government needs to finance an expenditure of G. It has three instruments available: labor income taxes, τ^h, capital taxes, τ_R^k,[8] and lump-sum transfers T (which can be positive or negative). Let w be the wage rate and R the gross interest rate.

Definition 1 *A tax-distorted competitive equilibrium is* $(K, h_L, h_H, w, R, \tau^h, \tau_R^k, T)$ *such that*

1. (K, h_L, h_H) *solves*

$$\max_{a,h_L,h_H} u(\omega - a, \bar{h}) + \beta E[u(c_i, h_i)], \quad s.t.\ c_i = (1-\tau^h)we_ih_i + (1-\tau_R^k)Ra + T;$$

2. $R = F_K(K,N)$, $w = F_N(K,N)$, *where* $N = \pi_L e_L h_L + \pi_H e_H h_H$;

3. *and,* $\tau^h wN + \tau_R^k RK = G + T$.

[8] Below we denote capital *income* taxes by τ^k, but here it is more convenient to use τ_R^k.

The Ramsey problem is to choose τ^h, τ_R^k, and T to maximize welfare in equilibrium. Since households are ex-ante identical there is no ambiguity about which welfare function to use. If there is no risk, i.e. $e_L = e_H$, the households are also ex-post identical and the usual representative-agent result applies: since lump-sum taxes are available, it is optimal to obtain all revenue via this undistortive instrument and set $\tau^h = \tau_R^k = 0$. When there is risk, this is no longer the case:[9]

Proposition 1 *The optimal tax system is such that*

$$\tau^h = \frac{\Omega}{1 - N + \gamma\Omega}, \quad \text{and} \quad \tau_R^k = \frac{(1-\gamma)\tau^h}{1-\gamma\tau^h},$$

where

$$\Omega \equiv \frac{\pi_L(1-e_L)u_{c,L} + \pi_H(1-e_H)u_{c,H}}{\pi_L u_{c,L} + \pi_H u_{c,H}} \geq 0.$$

The proof of this result and other claims made in this section can be found in Appendix B.[10] Notice that Ω can be interpreted as a measure of the planner's distaste for risk: it is zero if there is no risk and increases when risk is increased via a mean-preserving spread. Thus, it is evident from the formula for τ^h that labor income taxes are increasing in the amount of risk faced by households. This effectively provides insurance to households since it reduces the proportion of total household income that is risky. The optimal tax system, then, balances this provision of insurance with the reduction of distortions. Capital taxes do not affect the risk faced by households, but do allow the planner to mitigate some of the distortion caused by labor taxes via wealth effects: taxing capital reduces wealth in period 2 which increases labor supply.[11]

2.2 The effect of inequality

Consider the environment described above but replacing productivity risk with initial wealth inequality. That is, suppose that $e_L = e_H = 1$, and that the initial endowment can take two values: ω_L for a proportion p_L of households and $\omega_H > \omega_L$ for the rest. Let $\bar{\omega}$ denote the

[9]In a similar two-period environment, Gottardi et al. (2016) establish some properties of the solution to the Ramsey problem for general utility functions. They do, however, impose assumptions about the sign of general equilibrium effects, which are satisfied for the utility function considered here.

[10]Appendix B also discusses the case with both risk and inequality and connections with the results of Dávila, Hong, Krusell, and Ríos-Rull (2012) who study the related issue of constrained inefficiency in this environment.

[11]When there are no wealth effects on labor supply, a case considered in an earlier version of this paper, Dyrda and Pedroni (2016), optimal capital income taxes are set to zero.

average endowment. In this economy, the concept of optimality is no longer unambiguous. For the utilitarian welfare function we can show that:

Proposition 2 *If $\sigma = 1$,*[12] *then the optimal tax system is such that*

$$\tau_R^k = \frac{\gamma + \beta}{\beta} \frac{\Lambda}{\bar{\omega} - K + \Lambda}, \quad \textit{and} \quad \tau^h = 0,$$

where

$$\Lambda \equiv \frac{p_L(K - a_L)u_{c,L} + p_H(K - a_H)u_{c,H}}{p_L u_{c,L} + p_H u_{c,H}} \geq 0.$$

Here, Λ can be interpreted as a measure of the planner's distaste for inequality. The planner chooses a positive capital income tax which distorts savings decisions but allows for redistribution between households. The ex-ante wealth inequality is exogenously given. However, households with different wealth levels in period 1 save different amounts and have different asset levels in period 2. This endogenously generated asset inequality is the one the tax system is able to affect. A positive capital income tax directly reduces the proportion of households' income that depends on unequal asset income achieving the desired redistribution.

Optimal labor income taxes are set to zero. To see why, consider increasing labor taxation and rebating the extra revenue via a lump-sum. Since poorer households supply more labor, this change would have a negative redistributive effect. On the other hand, this would lead to higher savings for poor household which actually mitigates the distortion to their savings decisions. These effects exactly cancel each other.

The two-period example is useful for understanding some of the key trade-offs faced by the Ramsey planner, since it allows the levels of risk and inequality to be set exogenously. In the infinite horizon version of the SIM model, however, risk and inequality 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 labor income inequality and the distribution of assets over time. The asset level of a household in a particular period depends on the history of shocks the household has experienced. Therefore, capital income taxation affects both ex-ante and ex-post risk faced by households. Nevertheless, these results are useful for understanding some

[12] In the proof of this proposition, we obtain a more general result that applies for any σ. We impose this condition here to simplify the exposition, otherwise the formula for τ_R^k would be significantly more cumbersome, though it remains optimal to set $\tau^h = 0$.

features of the optimal fiscal policy in the infinite horizon model, as will become clear in what follows.

3 The Infinite-Horizon Model

In this model, time is discrete and infinite, indexed by t. There is a continuum of households with standard preferences $\mathbb{E}_0\left[\sum_t \beta^t u\left(c_t, h_t\right)\right]$ where c_t and h_t denote consumption and hours worked in period t. Individual labor productivity, denoted by $e \in E$ with $E \equiv \{e_1, \ldots, e_L\}$, follows a Markov process governed by the transition matrix Γ. Households can only accumulate a risk-free asset, a. Let the set of possible values for a be $A \equiv [\underline{a}, \infty)$, and let $S \equiv E \times A$. Households are indexed by the pair $(e, a) \in S$. Given a sequence of prices $\{r_t, w_t\}_{t=0}^{\infty}$, labor income taxes $\{\tau_t^h\}_{t=0}^{\infty}$, capital income taxes $\{\tau_t^k\}_{t=0}^{\infty}$, and lump-sum transfers $\{T_t\}_{t=0}^{\infty}$, each household at time t chooses $c_t\left(a, e\right)$, $h_t\left(a, e\right)$, and $a_{t+1}\left(a, e\right)$ to solve

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

subject to

$$(1 + \tau^c)c_t(a, e) + a_{t+1}(a, e) = \left(1 - \tau_t^h\right) w_t e h_t(a, e) + (1 + (1 - \tau_t^k)r_t)a + T_t$$

$$a_{t+1}(a, e) \geq \underline{a}.$$

Note that the value and policy functions are indexed by time, because policies $\{\tau_t^k, \tau_t^h, T_t\}_{t=0}^{\infty}$ and aggregate prices $\{r_t, w_t\}_{t=0}^{\infty}$ are time-varying. The consumption tax, τ^c, is a parameter.[13] 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 is a deterministic path for prices and for $\{\lambda_t\}_{t=0}^{\infty}$. It follows that 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, while δ is the depreciation rate. A representative firm exists that solves the usual static problem. The government needs to finance an exogenous constant stream

[13] 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 the federal government. We add τ^c as a parameter for calibration purposes.

of expenditure, G, and lump-sum transfers with taxes on consumption, labor income, and capital income. The government can also issue debt, $\{B_{t+1}\}$, subject to the constraint that the sequence is bounded. The government's intertemporal budget constraint is given by

$$G + r_t B_t = B_{t+1} - B_t + \tau^c C_t + \tau_t^h w_t N_t + \tau_t^k r_t (K_t + B_t) - T_t, \tag{3.1}$$

where C_t denotes aggregate consumption.

Definition 2 *Given K_0, B_0, $\{\tau_0^k, \tau_0^h, T_0\}$, an initial distribution λ_0, and a policy $\pi \equiv \{\tau_t^k, \tau_t^h, T_t\}_{t=1}^{\infty}$, a* ***competitive equilibrium*** *is a sequence of value functions $\{v_t\}_{t=0}^{\infty}$, an allocation $X \equiv \{c_t, h_t, a_{t+1},\ K_{t+1},\ N_t,\ B_{t+1}\}_{t=0}^{\infty}$, 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, $c_t(a,e)$, $h_t(a,e)$, 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 sequence of probability measures $\{\lambda_t\}_{t=1}^{\infty}$ satisfies*

$$\lambda_{t+1}(\mathcal{S}) = \int_{A\times E} Q_t\left((a,e), \mathcal{S}\right) d\lambda_t, \quad \forall \mathcal{S} \text{ in the Borel } \sigma\text{-algebra of } S,$$

   where Q_t is the transition probability measure;

4. *The government budget constraint, (3.1), holds and debt is bounded;*

5. *Markets clear,*

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

3.1 The Ramsey problem

We assume that, in period 0, the government announces and commits to a sequence of taxes $\{\tau_t^k, \tau_t^h, T_t\}_{t=0}^{\infty}$.

Definition 3 *Given K_0, B_0, and λ_0, for every policy π,* ***equilibrium allocation rules*** *$X(\pi)$ and* ***equilibrium price rules*** *$P(\pi)$ are such that $\{\pi, X(\pi), P(\pi)\}$ together with the corresponding $\{v_t\}_{t=0}^{\infty}$ and $\{\lambda_t\}_{t=1}^{\infty}$ constitute a competitive equilibrium. Given a welfare function $W(\pi)$, the* ***Ramsey problem*** *is to $\max_{\pi \in \Pi} W(\pi)$ subject to $X(\pi)$ and $P(\pi)$ being equilibrium allocation and price rules, and Π is the set of policies $\pi = \{\tau_t^k, \tau_t^h, T_t\}_{t=0}^{\infty}$ for which an equilibrium exists.*

In our benchmark experiments we assume that the Ramsey planner maximizes the utilitarian welfare function: the ex-ante expected lifetime utility of a "newborn" household who has its initial state, (a_0, e_0), chosen at random from the initial stationary distribution λ_0. The planner's objective is, thus, given by

$$W(\pi) = \int_S \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u\left(c_t\left(a_0, e_0 | \pi\right), h_t\left(a_0, e_0 | \pi\right)\right) \right] d\lambda_0.$$

We consider alternative welfare functions in Sections 6 and 9.

3.2 Solution method

Solving the Ramsey problem as stated would involve searching in the space of infinite sequences of fiscal instruments. To convert the problem into a finitely dimensional one we assume the existence of a Ramsey steady state—in the long run, all optimal fiscal instruments, including government debt, become constant and the economy settles in a final stationary equilibrium.[14] To decrease the dimensionality of the problem further, we build on Judd (2002) and parameterize the time paths of fiscal instruments as follows:

$$x_t = \left(\sum_{i=0}^{m_{x0}} \alpha_i^x P_i(t) \right) \exp\left(-\lambda^x t\right) + \left(1 - \exp\left(-\lambda^x t\right)\right) \left(\sum_{j=0}^{m_{xF}} \beta_j^x P_j(t) \right), \quad t \le t_F, \tag{3.2}$$

where x_t can be any of the fiscal instruments τ_t^k, τ_t^n, or T_t; $\{P_i(t)\}_{i=0}^{m_{x0}}$ and $\left\{P_j(t)\right\}_{j=0}^{m_{xF}}$ are families of Chebyshev polynomials, $\{\alpha_i^x\}_{i=0}^{m_{x0}}$ and $\left\{\beta_j^x\right\}_{j=0}^{m_{xF}}$ are weights on the consecutive elements of the family, λ^x controls the convergence rate of the fiscal instrument, and t_F is the period after which the instrument becomes constant. The orders of the polynomial approximations are given by m_{x0} and m_{xF} for the short-run and long-run dynamics. Given the calibrated initial stationary equilibrium, for any policy with instruments satisfying equation (3.2) we can compute the transition to the corresponding final stationary equilibrium,

[14] By stationary equilibrium we mean that all objects in Definition 2 become time invariant.

and evaluate welfare. We, then, pick the parameters that determine the policy to maximize welfare.

To implement this method we need to choose the orders of the Chebyshev polynomials. Generally, the larger they are the better the approximation is. In practice, however, as pointed out by Judd (2002), researchers should be interested in the smallest order that yields an acceptable approximation. Accordingly, we start with small orders and increase them for each instrument until the welfare gains from additional orders and changes in the instruments themselves are negligible. In our baseline experiment, we arrive at $m_{\tau_k 0} = m_{\tau_n 0} = 2$, $m_{\tau_k F} = m_{\tau_n F} = 0$, $m_{T0} = 2$ and $m_{TF} = 4$.[15] We set the terminal period at which taxes become constant to be $t_F = 100$,[16] and the upper bound on the capital income tax $\bar{\tau}_k = 100$ percent following the Ramsey literature.[17] Given these choices, we end up with the following 17 parameters:

$$\pi_A = \{\alpha_0^k, \alpha_1^k, \alpha_2^k, \beta_0^k, \lambda^k, \alpha_0^n, \alpha_1^n, \alpha_2^n, \beta_0^n, \lambda^n, \alpha_1^T, \alpha_2^T, \alpha_3^T, \alpha_4^T, \beta_0^T, \beta_1^T, \lambda^T\}, \tag{3.3}$$

which determine the time paths of fiscal instruments.

In order to solve problem described above we design a numerical algorithm for global optimization, based on insights from Guvenen (2011), Kan and Timmer (1987a), and Kan and Timmer (1987b). A detailed description is contained in the Appendix D.3, here we present a brief overview of the procedure. The algorithm is divided into two stages: a global and a local one. In the global stage we draw from a quasi-random sequence a very large number of policies in the domain of π_A. We compute transition and evaluate welfare $W(\pi_A)$ for each of those policies and select the ones that yield the highest levels of welfare. The selected policies are then clustered: 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, or until the bounds on parameters converge. Then, we pick the global optimum from the set of local optima.[18]

[15] In Appendix G.3 we discuss how the optimal policy changes as we gradually increase the number of choice variables.

[16] This is different from the length of the transition, which we set to 250 years so the economy has an additional 150 years to converge to a new stationary equilibrium. In Appendix G.4, we show that 100 is enough years of tax change. This can also be appreciated from the fact that all fiscal instruments stop moving well before this limit is reached.

[17] In Appendix K.5 we show how the policy is affected for different choices for $\bar{\tau}_k$, whereas in Appendix H we consider the case without any upper bound.

[18] The baseline experiment was conducted using 1200 cores on the Niagara supercomputer at the University

4 Calibration

We calibrate the initial stationary equilibrium of the model to replicate key properties of the US economy relevant for the shape of the optimal fiscal policy. We use three sets of statistics to discipline model parameters: (i) time series of macroeconomic data from 1995 to 2007, (ii) cross-sectional, distributional moments on hours worked, wealth, and earnings, and (iii) panel data on the dynamics of labor income. Even though it is understood that all model parameters impact all equilibrium objects, the discussion below associates some parameters to specific empirical targets for clarity of exposition. In total, we have 38 parameters in the model and we use 45 targets to discipline them, hence the system is overidentified. Parameter values, targeted statistics, and their model counterparts are presented in Tables 1 and 2. Appendix A contains a detailed description of how we calculated the targets from the data.

4.1 Households versus individuals

The unit of analysis in the model is a *household* rather than an individual. Thus, we consistently measure all the relevant statistics in the data at the household level using the equivalence scales proposed by the US Census. We then interpret consumption, hours, and asset positions in the household problem (3) in per-capita terms within the household.

4.2 Preferences and technology

The discount factor, β, is chosen to match a capital-output ratio of 2.5.[19] The two parameters in the balanced-growth-path utility function (2.1), γ and σ are disciplined with three targets: (1) an intertemporal elasticity of substitution (IES) of 0.65 (which implies a relative risk aversion of 1.55), (2) the average hours worked of employed households in the Current Population Survey (CPS) between 1995 and 2007, which is equal to 0.32, and (3) an average Frisch elasticity of labor supply of 0.6.

The model does a better job of matching targets overall for higher levels of IES. Though IES values from 0.5 to 1 are common in the quantitative public finance literature, most papers seem to prefer numbers closer to 0.5. So, balancing comparability with this literature and overall fit, we choose 0.65. Since household-level Frisch elasticities depend on the household's labor supply, we define the average Frisch elasticity as the sum of household-level Frisch elasticities weighted by the how much labor the household supplies relative to the aggregate,

of Toronto, see Ponce et al. (2019) and Appendix D.3 for details about the cluster.

[19] Capital is defined as nonresidential and residential private fixed assets and purchases of consumer durables.

that is

$$\Psi \equiv \int_{A\times E} \frac{h(a,e)}{H} \left(\frac{\gamma\sigma + (1-\gamma)}{\sigma} \frac{(1-h(a,e))}{h(a,e)} \right) d\lambda_0 \left(a,e\right). \tag{4.1}$$

This statistic is a measure of the intensive-margin response of household labor. The target of 0.6 incorporates higher flexibility within a household and the relatively higher elasticity of female labor supply.[20] We conduct sensitivity analysis with respect to our choices for the IES and Frisch elasticity in Section 9.

The production function, net of depreciation, is given by $f(K,N) = K^{\alpha}N^{1-\alpha} - \delta K$. The depreciation rate, δ, is set to match an investment-to-output ratio of 27 percent, and the capital share, α, to its empirical counterpart of 0.38. Finally, to discipline the household borrowing constraint, $\underline{a}$, we target the fraction of households with negative net worth in the 2007 Survey of Consumer Finances (SCF), which is 9.7 percent.

4.3 Fiscal policy

For the tax rates in the initial stationary equilibrium we use the effective average tax rates computed by Trabandt and Uhlig (2012) from 1995 to 2007. We set initial capital income tax to 41.5 percent, labor income tax to 22.5 percent and consumption tax to 4.7 percent. We discipline the lump-sum transfer, T, by targeting average transfer-to-GDP ratio in the US from 1995 to 2007, which amounts to 11.4 percent.[21] We set government debt to output ratio in the initial equilibrium to be 61.5 percent, averaging out federal debt over GDP from 1995 to 2007. The calibrated value implies a government expenditure to output ratio of 8.9 percent, while the data counterpart (federal government expenditure) for the relevant period is approximately 6.9 percent. Further, we also closely approximate the actual income tax schedule—see Figure 1.

4.4 Labor productivity process

The stochastic process for individual labor productivity levels, e, is calibrated to match statistical properties of the labor income process as well as the cross-sectional distributions

[20] To check whether the extensive-margin of labor supply is also in line with the data, we consider the transitional dynamics following a temporary 1 percent increase in the wage rate and compute the elasticity of employment with respect to this change. Aggregate hours, H, can be expressed as $H = m \times h$, where m denotes the employment rate and h mean working hours. It follows that the corresponding elasticities satisfy $\eta_H = \eta_m + \eta_h$. Our calibration implies that, on impact, $\eta_m = 0.57$ and $\eta_h = 0.45$. The contribution of the extensive margin is in line with the findings in Erosa, Fuster, and Kambourov (2016).

[21] We define transfers in the data as personal current transfer receipts, which include social security transfers, medicare, medicaid, unemployment benefits, and veteran benefits. We choose this for two reasons: First, we include retired and unemployed households in our inequality moments. Second, lump-sum transfers in the model can be interpreted as a basic income in the case of not working.

Table 1: Benchmark Model Parameters

Description	Parameter	Value
Preferences and technology		
Consumption share	γ	0.510
Preference curvature	σ	2.069
Discount factor	β	0.954
Capital share	α	0.378*
Depreciation rate	δ	0.104
Borrowing constraint	$\underline{a}$	−0.078
Fiscal policy		
Capital income tax (%)	τ^k	41.5*
Labor income tax (%)	τ^n	22.5*
Consumption tax (%)	τ^c	4.7*
Government expenditure	G	0.069
Transfers	T	0.088
Labor productivity process		
Productivity process curvature	η	1.153

Persistent shock

$$\Gamma_P = \begin{bmatrix} 0.994 & 0.002 & 0.004 & 3E{-}5 \\ 0.019 & 0.979 & 0.001 & 9E{-}5 \\ 0.023 & 0.000 & 0.977 & 5E{-}5 \\ 0.000 & 0.000 & 0.012 & 0.987 \end{bmatrix} \quad e_P = \begin{bmatrix} 0.580 \\ 1.153 \\ 1.926 \\ 27.223 \end{bmatrix}$$

Transitory shock

$$P_T = \begin{bmatrix} 0.263 \\ 0.003 \\ 0.556 \\ 0.001 \\ 0.001 \\ 0.176 \end{bmatrix} \quad e_T = \begin{bmatrix} -0.574 \\ -0.232 \\ 0.114 \\ 0.133 \\ 0.817 \\ 1.245 \end{bmatrix}$$

Note: $*$ denotes exogenously set parameters.

of hours worked, wealth, and earnings. We model it as a product 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. We normalize the average productivity to one, so we are left with 26 parameters in the labor income process to choose.

There are two approaches commonly used in the literature. The first is to reduce the number of parameters using a discretization procedure, such as Tauchen (1986) or Rouwen-

[22]In the notation of the model, $\Gamma = \Gamma_P \otimes \mathrm{diag}(P_T)$, and $e = e_P + e_T e_P^{\eta}$. For instance, if $\eta = 0$, the transitory shocks are additive, whereas, if $\eta = 1$, they are multiplicative.

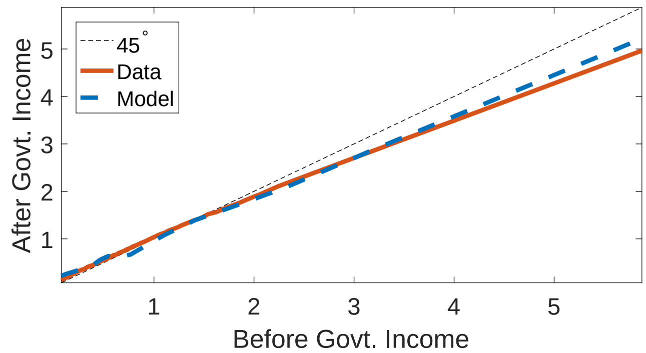


Figure 1: Income tax schedule

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 corresponding mean.

horst (1995), and target a small set of moments usually only focusing on the labor-income process itself. The second approach, put forward by Castañeda et al. (2003), abstracts from labor-income process targets and, instead, targets enough distributional moments to identify the large set of parameters. We largely follow this second approach but, importantly, we also include moments of the labor income process itself. This gives us, at the same time, the ability to match important inequality measures *and* moments of the labor income process, including higher moments such as skewness and kurtosis which the first method struggles with. Thus, we discipline the amount of inequality and risk that households face.

Inequality. We target the share owned by every quintile, the Gini coefficient, and the share owned by the bottom and top 5 percent of the wealth, earnings, and hours distributions. For wealth and earnings we use data from the SCF, and for hours we use the Current Population Survey (CPS). To account for the joint distribution of earnings and wealth we also target the cross-sectional correlation between them. Lastly, we target the fraction of employed households from the CPS data.

Risk. Pruitt and Turner (2020) document statistical properties of the labor income process for households using administrative data from the IRS. We exploit their findings and compute the variance, Kelly skewness, and Moors kurtosis of the growth rates of labor income, which we target. These moments, however, do not include self-employed households. To deal with this, we identify one element of the vector e_P with self-employed status. We think of this state as representing, in a reduced form, entrepreneurial opportunities of households in our model. Entrepreneurs, on average, earn higher incomes and account for a disproportional fraction of wealth in the SCF data which we include as targets. On the other hand, for consistency, we exclude households in this state from the computation of the labor-income moments.[23]

4.5 Model performance

Table 3 presents income sources over the quintiles of income. The composition of income, especially of the consumption-poor households, plays an important role in determining the optimal fiscal policy. The fraction of uncertain labor income determines the strength of the insurance motive while the fraction of unequal asset income affects the redistributive motive. Our calibration delivers, without targeting, a good approximation of the composition of household income. Figures 3a–3d present the model's fit with the cross-sectional distributions of the targeted wealth, earnings, and hours. The last two panels of the figure show that the model also approximates well the untargeted distributions of income and consumption.

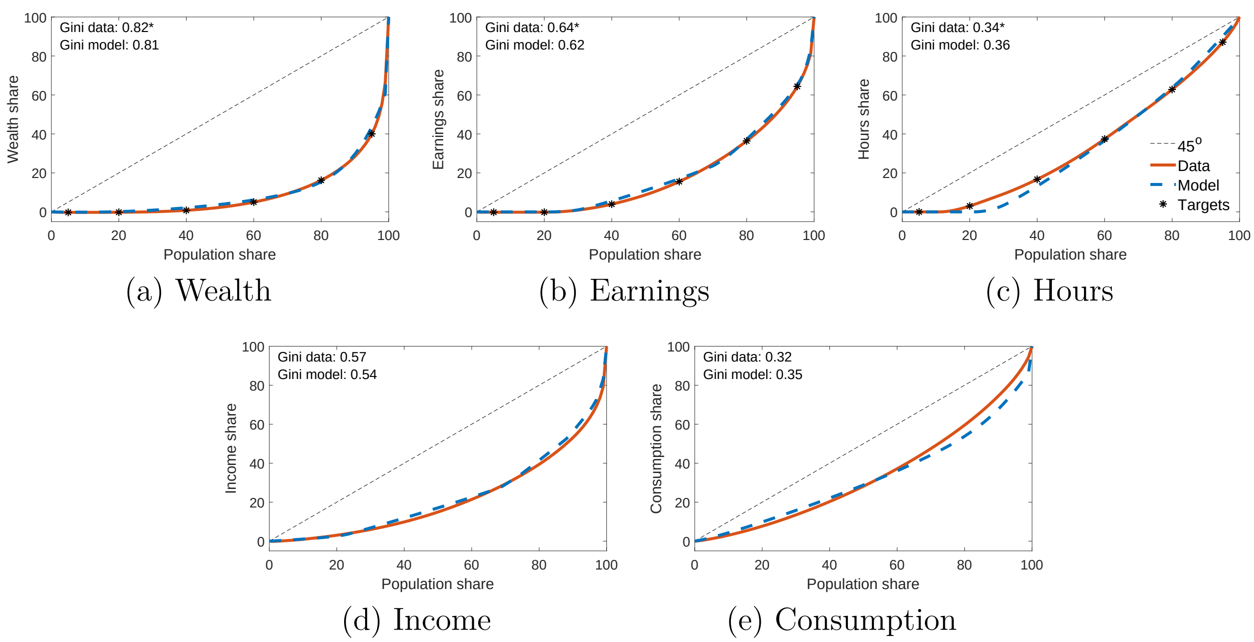


Figure 2: Fit to Inequality Data

5 Main Results

The optimal paths for the fiscal policy instruments are presented in Figure 3. The capital income tax is front-loaded, hitting the upper bound for 16 years, and decreasing to 26 percent in the long run. The labor income tax drops on impact to 9 percent and then monotonically

[23]A similar strategy has been employed by Kindermann and Krueger (2014) and Nakajima and Ríos-Rull (2019).

Table 2: Benchmark Model Economy: Target Statistics and Model Counterparts

(1) Macroeconomic aggregates

	Target	Model
Intertemporal elasticity of substitution	0.65	0.65
Average Frisch elasticity (Ψ)	0.60	0.62
Average hours worked	0.32	0.33
Capital to output	2.50	2.49
Capital income share	0.38	0.38
Investment to output	0.26	0.26
Transfer to output (%)	11.4	11.4
Debt to output (%)	61.5	61.5
Share of workers (%)	79.0	79.3
Fraction of hhs with negative net worth (%)	9.7	7.9
Correlation between earnings and wealth	0.43	0.43

(2) Cross-sectional distributions

	Bottom (%)	Quintiles					Top (%)	Gini
	0–5	1st	2nd	3rd	4th	5th	95–100	
Wealth								
US data	−0.2	−0.2	1.0	4.2	11.2	83.8	60.0	0.82
Model	−0.1	0.1	2.0	4.0	9.3	84.5	56.4	0.81
Earnings								
US data	−0.2	−0.2	4.1	11.6	20.9	63.6	35.6	0.64
Model	0.0	0.0	5.7	11.3	20.2	62.8	34.8	0.62
Hours								
US data	0.0	3.0	13.7	20.7	25.4	37.2	12.9	0.34
Model	0.0	0.0	13.2	23.4	27.1	36.3	10.0	0.36

(3) Statistical properties of labor income

	Target	Model
Variance of 1-year growth rate	2.33	2.32
Kelly skewness of 1-year growth rate	−0.12	−0.13
Moors kurtosis of 1-year growth rate	2.65	2.28

(4) Self-employed status statistics

	Target	Model
Share in population (%)	12.5	12.7
Share of wealth (%)	45.8	38.9
Share of earnings (%)	28.7	30.5

Table 3: Income Sources of Households by Quintile of Income

Quintile	Model			US data		
	Labor	Asset	Transfer	Labor	Asset	Transfer
1st	80.1	0.2	19.7	83.6	0.4	16.1
2nd	77.0	2.6	20.4	86.5	1.1	12.3
3rd	74.1	5.3	20.5	85.6	1.9	12.5
4th	74.8	9.4	15.8	84.1	3.8	12.2
5th	63.1	31.2	5.7	70.4	21.4	8.2
All	70.4	16.7	12.9	77.3	12.3	10.4

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 Consumer Finances.

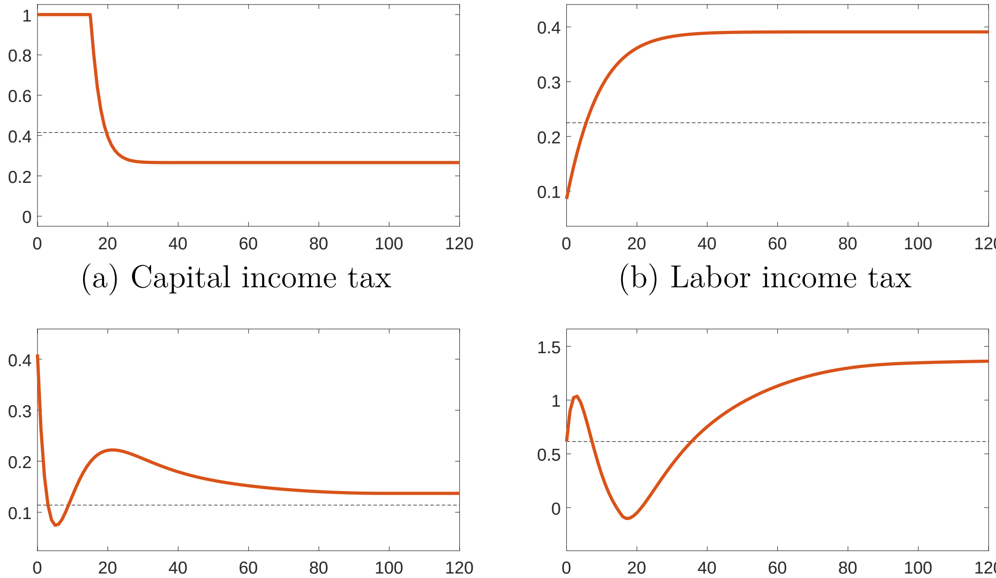
increases to 39 percent in the long run. Lump-sum transfers jump to 40 percent of GDP on impact, follow a U-shaped pattern in the short-run and, starting from period 22, fall monotonically toward 15 percent of GDP in the long run. The government-debt-to-output rises in the initial periods. Then, since the capital income is kept at the upper bound but transfers fall, the government accumulates assets. Finally, the reduction of capital income tax combined with the increase in transfers leads to an increase in government debt toward 154 percent of GDP in the long run. This policy yields welfare gains equivalent to a 3.52 percent permanent increase in the consumption of all households. Figures 4 and 5 summarize the main effects of the optimal policy on aggregates and on the distribution of households.[24]

In what follows, we briefly describe aggregate and distributional statistics that summarize the effects of the Ramsey policy. Then, to understand the economic forces behind the results and to inspect the role played by each fiscal instrument, we introduce a decomposition of the welfare effects and conduct policy perturbations around the optimum.

5.1 Aggregates

High capital income taxes in the initial periods lead to a reduction in the capital stock of about 10 percent. The substantial fall in these taxes later on does not imply a recovery for three reasons: (1) government debt increases, which crowds out private capital, (2) labor decreases over time as a result of higher labor income taxes, which reduces the marginal

[24] Appendix K.1 contains a more exhaustive list of figures.



(a) Capital income tax (b) Labor income tax

(c) Lump-sum/Initial output (d) Debt/Initial output

Figure 3: Optimal Fiscal Policy: Benchmark

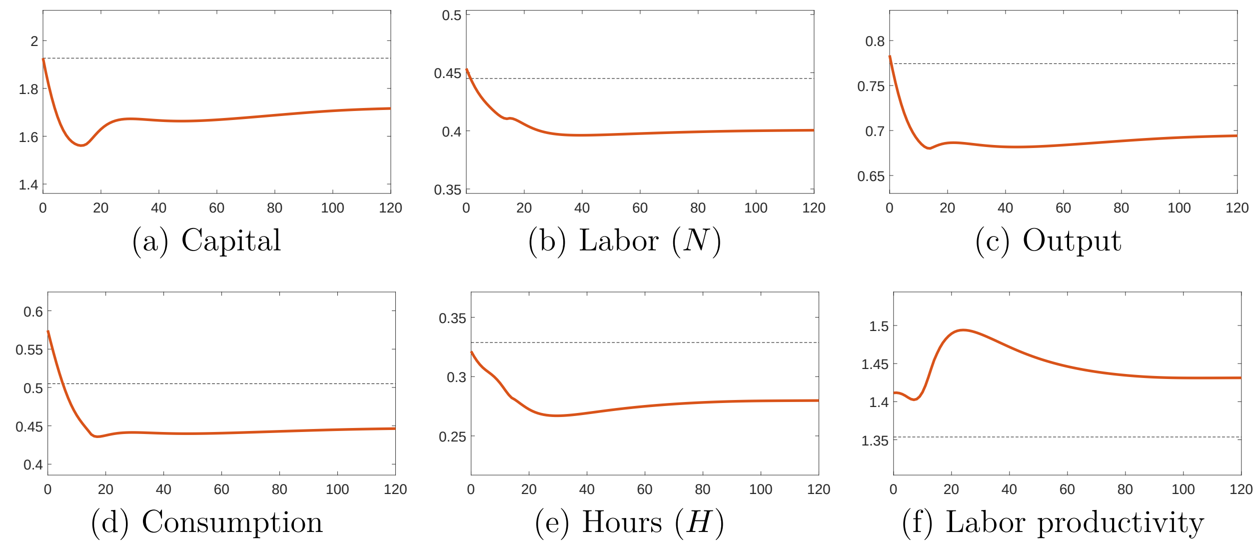
Note: Black dashed lines: initial stationary equilibrium; Red solid curves: optimal transition.

product of capital, and (3) the optimal policy implies a reduction in risk faced by households, which reduces precautionary savings.

Aggregate consumption increases on impact, then decreases towards a level also about 10 percent lower than the pre-policy-change value. The low after-tax interest rates account for the downward slope in the initial periods, and the long-run decrease is consistent with the decrease in output associated with the overall lower long-run levels of capital and labor.

Even with lower labor income taxes in the initial periods, aggregate hours fall on impact. This is due to the redistribution achieved by the increase in initial capital income taxes and lump-sum transfers. The associated wealth effects on labor supply reduce the labor supply of the more numerous lower-productivity households. The subsequent reduction in hours worked are due to increasing labor income taxes. In the long run, aggregate hours fall by 15 percent relative to the initial equilibrium.

Most of the welfare gains associated with this policy come from redistribution and insurance. However, the average household is also better off under this reform—see Section 5.3. This is partially due to the higher levels of leisure associated with the reduction in hours worked. More importantly, though, it is due to a *more efficient allocation of labor supply*. The redistribution achieved by the policy makes low-productivity households relatively wealthier, and the associated wealth effects reduce their labor supply.[25] The opposite occurs with high-productivity households. These changes result in a significant increase in average labor productivity—measured by the ratio of effective labor to hours worked—which can be seen in Figure 4f. In Section 6, we show that, as a result of this mechanism, even a planner that does not value reductions in inequality would be in favor of some amount of redistribution.



(a) Capital (b) Labor (N) (c) Output

(d) Consumption (e) Hours (H) (f) Labor productivity

Figure 4: Optimal Fiscal Policy: Aggregates

Note: Black dashed lines: initial stationary equilibrium; Red solid curves: optimal transition.

5.2 Distributional effects

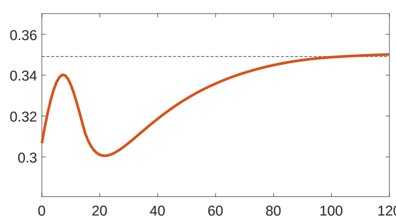
The optimal policy implies a reduction in the amount of inequality and risk faced by households. This is achieved, to a large extent, simply by the increase in the share of households' income that comes from equal and certain lump-sum transfers, which we illustrate in Figure 5e. This translates into less overall risk and inequality. To show this in a compact way it is useful to define a consumption–hours composite, $c^{\gamma}(1-h)^{1-\gamma}$, which is the term that enters the household period utility function. In Figures 5c and 5d, we show that the optimal policy implies a reduction in the amount of inequality (measured by the Gini coefficient of the composite) and risk (measured by the variance of the growth rate of the composite) that households face.

The reduction in inequality of the composite, however, masks a different effect of the policy

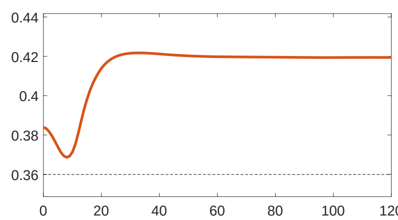
[25] Marcet, Obiols-Homs, and Weil (2007) show that wealth effects on labor supply also play an important role in determining whether there is over- or under-accumulation of capital in the SIM model.

on consumption and hours. Figures 5a and 5b show that the policy implies a significant reduction in consumption inequality, but an increase in hours inequality. This increase comes from the more efficient allocation of labor supply highlighted above.

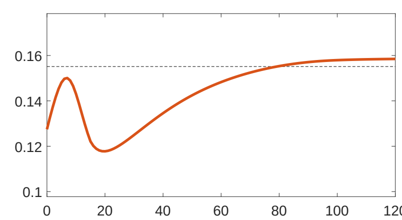
Figure 5f shows how the welfare gains are distributed between households with different levels of wealth and labor income. In line with the redistribution achieved, to a large extent via high initial capital income taxes, wealthy households lose and asset-poor households win. Conditional on wealth quantile, however, the welfare gains remain similar across quantiles of labor income. This is because the provision of insurance benefits all households in a similar way—risk is more consequential to low-productivity households, but since transitory shocks are roughly multiplicative these households actually face less income risk.



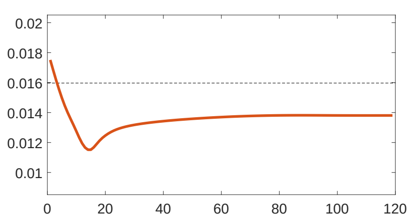
(a) Consumption Gini



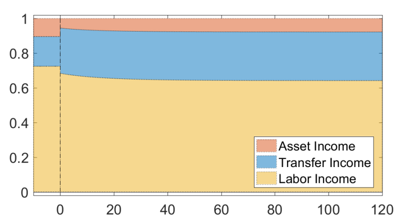
(b) Hours Gini



(c) Gini of $c^{\gamma}(1-h)^{1-\gamma}$



(d) Var.of $c^{\gamma}(1-h)^{1-\gamma}$ growth



(e) Income shares

(f) Conditional welfare

Figure 5: Optimal Fiscal Policy: Distributional Effects

Notes: Panels (a)–(d): Black dashed lines: initial stationary equilibrium; Red solid curves: optimal transition. Panel (f): In the axis we have 20-quantiles of wealth and labor income. The size of the tiles is proportional to the density of households in the initial stationary distribution. The color of the tile represents the welfare gain or loss associated with the optimal policy conditional on the household's level of wealth and labor income.

5.3 Sources of welfare improvement

In this section, we present a decomposition of average welfare gains that is helpful for understanding the properties of the optimal fiscal policy. This decomposition is similar to the ones introduced by Benabou (2002) and Floden (2001), but here we allow not only for welfare comparisons between steady states, but also for transitional effects of policy.

Average welfare gains. Consider a policy reform and denote by $\{c_t^j, h_t^j\}$ the equilibrium consumption and labor paths of a household with and without the reform, with $j = R$ or $j = NR$ respectively. The average welfare gain, Δ, that results from implementing the reform is defined as the constant (over time and across households) percentage increase to c_t^{NR} that equalizes the utilitarian welfare to the value associated with the reform; that is,

$$\int \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u \left((1+\Delta) c_t^{NR}, h_t^{NR} \right) \right] d\lambda_0 = \int \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u \left(c_t^R, h_t^R \right) \right] d\lambda_0, \tag{5.1}$$

where λ_0 is the initial distribution over states (a_0, e_0). These welfare gains associated with the utilitarian welfare function can be decomposed into three effects which we introduce one at a time.

1. Level effect. First, the average welfare gain can come from increases in the utility of the average household. Reductions in distortive taxes or a more efficient allocation of resources achieve this goal. This is the only relevant effect in a representative agent economy without any source of heterogeneity. Let the aggregate level of c_t and h_t at each t be

$$C_t^j \equiv \int c_t^j d\lambda_t^j, \quad \text{and} \quad H_t^j \equiv \int h_t^j d\lambda_t^j,$$

where λ_t^j is the distribution over (a_0, e^t) conditional on whether or not the reform is implemented—e^t denotes a history of productivity realizations from period 0 to t. The level effect, Δ_L, is then given by

$$\sum_{t=0}^{\infty} \beta^t u \left((1+\Delta_L) C_t^{NR}, H_t^{NR} \right) = \sum_{t=0}^{\infty} \beta^t u \left(C_t^R, H_t^R \right). \tag{5.2}$$

2. Insurance effect. Since households are risk averse, average welfare increases if, conditional on a household's initial asset and productivity state, the riskiness of its future consumption and labor paths is reduced. A tax reform that transfers from the ex-post lucky to the ex-post unlucky reduces the risk faced by households. To define this component precisely, first let $\{\bar{c}_t^j(a_0, e_0), \bar{h}_t^j(a_0, e_0)\}$ denote a certainty-equivalent sequence of consumption and labor conditional on a household's initial state that satisfies

$$\sum_{t=0}^{\infty} \beta^t u \left(\bar{c}_t^j(a_0, e_0), \bar{h}_t^j(a_0, e_0) \right) = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u \left(c_t^j, h_t^j \right) \right]. \tag{5.3}$$

Next, let $\bar{C}_t^j$ and $\bar{H}_t^j$ denote aggregate certainty equivalents, that is

$$\bar{C}_t^j = \int \bar{c}_t^j(a_0, e_0) d\lambda_0, \quad \text{and} \quad \bar{H}_t^j = \int \bar{h}_t^j(a_0, e_0) d\lambda_0, \quad \text{for } j = R, NR. \tag{5.4}$$

The insurance effect, Δ_I, is defined by

$$1 + \Delta_I \equiv \frac{1 - p_{risk}^R}{1 - p_{risk}^{NR}}, \quad \text{where } \sum_{t=0}^{\infty} \beta^t u\left((1 - p_{risk}^j) C_t^j, H_t^j\right) = \sum_{t=0}^{\infty} \beta^t u\left(\bar{C}_t^j, \bar{H}_t^j\right). \tag{5.5}$$

Here, p_{risk}^j is the welfare cost of risk in the economies with and without reform.

3. Redistribution effect. The utilitarian welfare gain increases if inequality across households with different initial states is reduced. A tax reform reduces the inequality if it redistributes from rich (ex-ante lucky) to poor (ex-ante unlucky) households, that is by reducing the behind-the-veil-of-ignorance risk. Formally, the redistribution effect, Δ_R, can be defined as

$$1 + \Delta_R \equiv \frac{1 - p_{ineq}^R}{1 - p_{ineq}^{NR}}, \quad \text{where } \sum_{t=0}^{\infty} \beta^t u\left((1 - p_{ineq}^j) \bar{C}_t^j, \bar{H}_t^j\right) = \int \sum_{t=0}^{\infty} \beta^t u\left(\bar{c}_t^j(a_0, e_0), \bar{h}_t^j(a_0, e_0)\right) d\lambda_0. \tag{5.6}$$

Analogously to p_{risk}^j, p_{ineq}^j denotes the cost of inequality. Redistribution, according to this definition, is also a type of insurance but with respect to the ex-ante risk a household faces concerning which initial condition (a_0, e_0) they will receive.

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

Proposition 3 *If preferences are such that, for any scalar x, $u(xc, h) = g(x)u(c, h)$ for some totally multiplicative function $g(\cdot)$, then*

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

Note that none of the elements of the decomposition are defined residually, hence this is indeed a decomposition and not a definition. Moreover, the balanced-growth-path utility satisfies the condition with $g(x) = x^{\gamma(1-\sigma)}$.

Choice of certainty equivalents. Notice that there can be many certainty-equivalent paths that satisfy equation (5.3). These paths could differ over time and over levels of consumption and labor. In general, these choices can affect the components of the decomposition. If the certainty equivalents for consumption and leisure follow parallel paths over time, however, these choices are immaterial.

Assumption 1 *The certainty equivalents display parallel patterns if* $\bar{c}_t^j(a_0, e_0) = \eta^j(a_0, e_0)\tilde{C}_t^j$, *and* $1 - \bar{h}_t^j(a_0, e_0) = \eta^j(a_0, e_0)(1 - \tilde{H}_t^j)$, *for some function* $\eta^j(a_0, e_0)$ *and paths* $\{\tilde{C}_t^j\}$, *and* $\{\tilde{H}_t^j\}$.

There are two ways in which this assumption is restrictive. First, it assumes that the certainty equivalents of households with different initial conditions are a proportion of the same paths, with only the degree of proportionality, $\eta^j(a_0, e_0)$, changing; this is the property we are referring to as "parallel patterns". Second, it assumes that the degree of proportionality applies in the same way to the path of consumption and leisure. Reasonable deviations from the first restriction lead to small changes in the results benchmark results in Table 4 below. The second restriction is more consequential, because the way one decomposes the differences between consumption and leisure affects the amount of curvature that is absorbed by the insurance and redistribution effects.[26] The choice of certainty equivalents, however, never matters for the magnitude of the level effect. Under this assumption, we can establish the following proposition.

Proposition 4 *For balanced-growth-path preferences, as specified in equation (2.1), if the certainty equivalents satisfy Assumption 1, then the components* Δ_L, Δ_I, *and* Δ_R *are independent of the paths* $\{\tilde{C}_t^j\}$, *and* $\{\tilde{H}_t^j\}$.

In particular, one could impose that the certainty-equivalent paths should follow their corresponding aggregates, that is $\tilde{C}_t^j = C_t^j$ and $\tilde{H}_t^j = H_t^j$. In any case, as long as Assumption 1 is satisfied this choice does not matter. Table 4 shows the welfare-decomposition results for our benchmark experiment and for experiments in which we hold each instrument fixed, one at a time.

Fixed instruments. Fixing capital income taxes at their initial steady-state level, while all other instruments are set to their optimal paths from the benchmark experiment, leads to

[26] More precisely, the degree of proportionality is taken to a different power if it multiplies only consumption, for instance, $(\eta c)^\gamma (1-h)^{1-\gamma} = \eta^\gamma (c)^\gamma (1-h)^{1-\gamma}$ versus if it multiplies consumption and leisure as in Assumption 1, $(\eta c)^\gamma (\eta(1-h))^{1-\gamma} = \eta (c)^\gamma (1-h)^{1-\gamma}$.

Table 4: Welfare Decomposition for the Benchmark Experiment and Fixed-Instrument Experiments

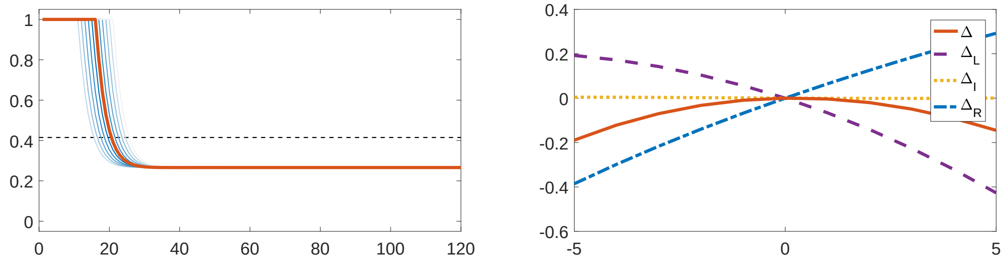
	Δ	Δ_L	Δ_I	Δ_R
Benchmark	3.52	0.23	1.19	2.07
Fixed capital income tax	0.77	−0.59	1.25	0.11
Fixed labor income tax	1.98	0.60	−0.34	1.71
Constant lump-sum	3.32	−0.13	1.31	2.12
Fixed debt-to-output	3.17	−0.22	1.35	2.02
Fixed prices	4.89	1.53	1.07	2.21

a substantial reduction in redistributive gains. The drop in the level effect is also significant, due to (1) the significantly higher long-run capital income taxes fixed at the initial level and (2) the loss of the productivity improvements that result from redistribution via wealth effects on labor supply. The second most welfare-relevant instrument is the labor income tax. Fixing it at its pre-reform level reduces average welfare by roughly 1.5 percentage points, mostly through the insurance channel. The time paths of both lump-sum transfers and government debt contribute marginally to average welfare with the losses coming mostly from the level effect. Both instruments are used mostly to smooth distortions over time. Finally, we consider an out-of-equilibrium experiment in which prices are kept at their initial values. Since aggregate output decreases in the benchmark results, with fixed prices households have relatively more income, which explains the larger level effect. The general equilibrium price effects on insurance and redistribution are relatively small: prices do not move much since capital and labor move together.

5.4 Perturbations around the optimal taxes

In this section we vary the taxes around the optimal paths and calculate the welfare decomposition at each step in order to better understand the main determinants of the optimal values. Notably, we consistently find that welfare peaks at the paths we found and falls in their neighborhood, which is indicative of the fact that we have indeed found an optimum. For each experiment, the entire path of lump-sum taxes is 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 16 years of taxes at the upper bound of 100 percent. Figure

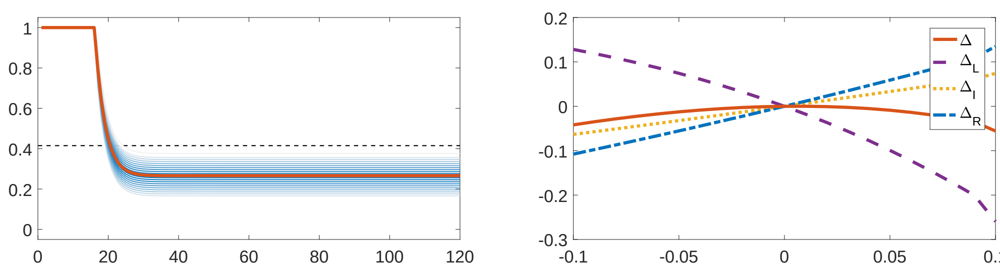


(a) Capital income tax (b) Welfare decomposition

Figure 6: Varying the Number of Years Capital Income Taxes are Kept at the Upper Bound

Note (a) Black dashed line: initial stationary equilibrium; Red solid and blue curves: optimal transition and perturbations of it; (b) the x-axis represents the movement in number of periods capital income taxes are kept in the upper bound from the optimum, y-axis shows change in the welfare gains in percent points.

6 shows what happens to the components of welfare if capital income taxes are kept at the upper bound for more or fewer periods. The effect on insurance is of second order and, in line with the result in Proposition 2, the relevant trade-off is between extra redistribution and negative distortionary effects. These two effects, however, largely offset each other, leading to a relatively flat average welfare function.

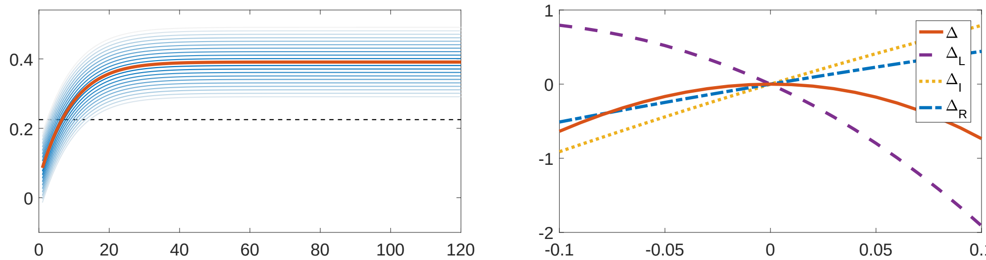


(a) Capital income tax (b) Welfare decomposition

Figure 7: Varying Long-Run Capital Income Taxes

Note (a) Black dashed line: initial stationary equilibrium; Red solid and blue curves: optimal transition and perturbations of it; (b) The x-axis represents the movement of long-run capital income taxes away from the optimum, y-axis shows change in the welfare gains in percent points.

Long-run capital income taxes. Varying the level of long-run capital income taxes yields the results in Figure 7. The changes considered here affect the path of capital income taxes starting in period 16, and therefore still have a sizable effect of ex-ante risk captured by the redistribution effect. The main difference relative to Figure 6 is that the insurance effect is of comparable magnitude to redistribution. As highlighted by Chamley (2001) and Acikgoz et al. (2018), far enough in the future every household's dependence on their initial condition fully dissipates, so that changes in income taxes have no effect on redistribution, but only on level and insurance. Indeed, in Section 6 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, notice again how flat the average welfare function is in response to relatively sizable changes in the path of capital income taxes.

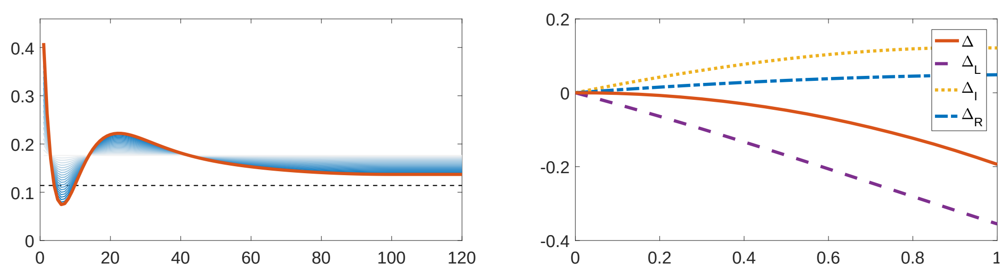


(a) Labor income tax (b) Welfare decomposition

Figure 8: Varying Labor Income Taxes

Note (a) Black dashed line: initial stationary equilibrium; Red solid and blue curves: optimal transition and perturbations of it; (b) The x-axis represents the movement of labor income taxes away from the optimum, y-axis shows change in the welfare gains in percent points.

Labor income taxes. Here we change the average level of labor income taxes up and down by 10 percentage points, leading to the results in Figure 8. First notice that the effect of changes in labor income taxes are an order of magnitude higher than the previous ones. Besides this quantitative difference, the main qualitative difference is that the insurance effect is larger than the redistribution effect. Hence, though labor income taxes do have important effects on ex-ante risk, the mechanism highlighted in Proposition 1 plays a more important role here. That is, a higher labor income tax which is rebated via lump-sum transfers (exactly the experiment here) effectively reduces the labor income risk to which households are exposed.



(a) Lump-sum/Initial output (b) Welfare decomposition

Figure 9: Varying Lump-Sum Transfers

Note (a) Black dashed line: initial stationary equilibrium; Red solid and blue lines: optimal transition and perturbations of it; (b) The x-axis represents the homotopy parameter between the initial optimal path at $x = 0$ and a flat path at $x = 1$, y-axis shows change in the welfare gains in percent points.

The path of lump-sum transfers. Figure 9 shows what happens to welfare when the path of lump-sum transfers is gradually replaced by a constant. This change leads a reduction in average welfare gains of about 0.2 percent. For households close enough to their borrowing constraints, the initial sharp front-loading of lump-sum transfers mitigates the distortions associated with high capital income taxes. Hence, moving to a flatter lump-sum path reduces the gains that occur via the level effect. It is also relevant to notice that, absent borrowing constraints, households would be indifferent to the timing of lump-sum transfers.[27] Since households do face borrowing constraints, however, they would, *ceteris paribus*, always prefer lump-sum transfers to be front-loaded as much as possible. The reason this is not optimal, and why lump-sum transfers actually increase in the medium run, is because front-loading lump-sum transfers to this extent would lead to a severe increase in government debt. The corresponding crowding out of capital would compound with the reduction that already occurs due to high initial capital income taxes and the reduction in precautionary savings that results from the extra insurance.

5.5 Long-run optimality conditions

Aiyagari (1995) analyzes optimal long-run capital income taxes in an environment similar ours.[28] He argues that the Ramsey planner's decision to move aggregate resources across time is risk-free and the associated Euler equation, in the long run, implies the modified golden rule. Lining this up with households' precautionary motivation for savings rationalizes positive long-run capital income taxes. Figure 10 shows that the modified golden rule is satisfied in our benchmark results. We view this as corroborating evidence for the accuracy of our numerical long-run results.

Acikgoz et al. (2018) have made advances in obtaining a better characterization of the long-run optimal tax system in the same environment as ours, except that they use a separable utility function. They argue that the long-run optimal tax system is independent of initial conditions and of the transition into it, and show that the modified golden rule and three additional optimality conditions must hold. In Appendix I, we extend their results to the BGP preferences used in this paper and show that our long-run results do satisfy those three additional conditions. We also compute the optimal paths using our method but with their calibration, and find long-run results that are consistent with their findings.

[27] Without borrowing constraints, the households' lifetime budget constraint would not be affected by a revenue-neutral change in the timing of lump-sum transfers (holding other taxes fixed). So, for this type of variation the Ricardian equivalence would hold. If instead we were considering a change in the timing of capital or labor income taxes, this would affect the risk faced by households, which would then violate Ricardian equivalence as in Barsky et al. (1986). Bhandari et al. (2017) formalize a similar argument.

[28] In Aiyagari's environment the planner does not have lump-sum taxes as an instrument, and chooses the level of government expenditure every period (which enters separately into households' utility function).

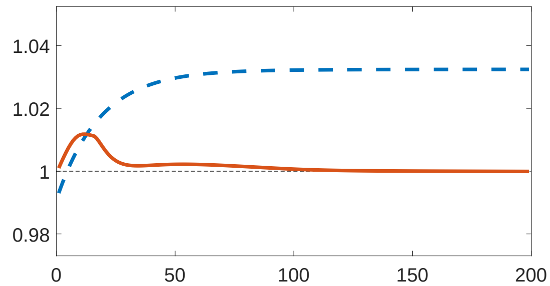


Figure 10: $\beta(1 + f_K(K,N))$

Note: Red solid curve: benchmark experiment; Dashed blue curve: optimal transition with constant policy (see Section 7.2).

Quantitative differences between our results and theirs must, therefore, be due to differences in the calibration and not the solution method.[29] In Appendix I we also compare the two calibrations and discuss in detail the likely roots of these differences. We also provide there an extensive discussion of the advantages and disadvantages of both numerical methods.

6 Maximizing Efficiency: The Role of Redistribution

The utilitarian welfare function, which we consider in our benchmark results, places equal Pareto weights on every household. This implies a particular social preference with respect to the equality-versus-efficiency trade-off. Here, we consider a different welfare function that rationalizes different preferences about this trade-off,

$$W^{\hat{\sigma}} = \left(\int \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t, h_t) \right]^{\frac{1-\hat{\sigma}}{1-\sigma}} d\lambda_0 \right)^{\frac{1-\sigma}{1-\hat{\sigma}}},$$

where λ_0 is the initial distribution over individual states (a_0, e_0). Following Benabou (2002), we refer to $\hat{\sigma}$ as the planner's degree of inequality aversion. If $\hat{\sigma} = \sigma$, maximizing W^{σ} is equivalent to maximizing the utilitarian welfare function. If $\hat{\sigma} \to \infty$, this becomes the Rawlsian welfare function. Finally, if $\hat{\sigma} = 0$, then maximizing W^0 is equivalent to maximizing efficiency. We formalize claim in the following proposition.

Proposition 5 *If the certainty equivalents satisfy Assumption 1, then, maximizing W^0 is equivalent to maximizing $(1 + \Delta_L)(1 + \Delta_I)$.*

[29]The most stark differences are that they find substantially higher optimal labor income taxes and debt-to-output ratios than we do. The higher levels of labor income taxes result, to a large extent, from stronger wealth effects on labor supply under their calibration. Their optimal policy yields very few borrowing-constrained households in the long run, making long-run debt even less consequential than in our economy.

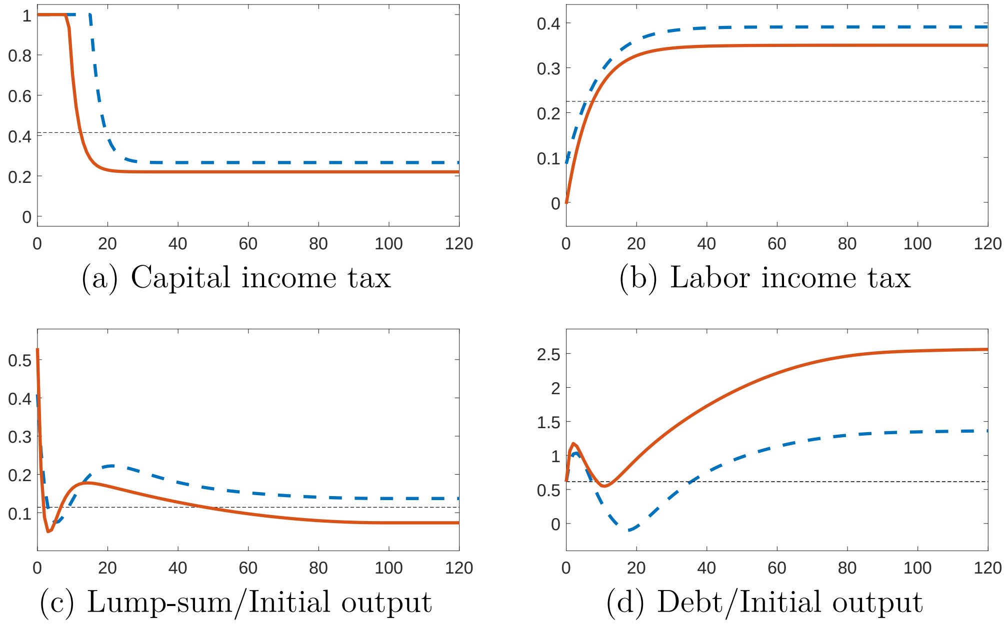


Figure 11: Optimal Fiscal Policy: Maximizing Efficiency

Note: Black dashed line: initial stationary equilibrium; Red solid curve: path that maximizes efficiency optimal transition; Blue dashed curve: path that maximizes the utilitarian welfare function (benchmark results).

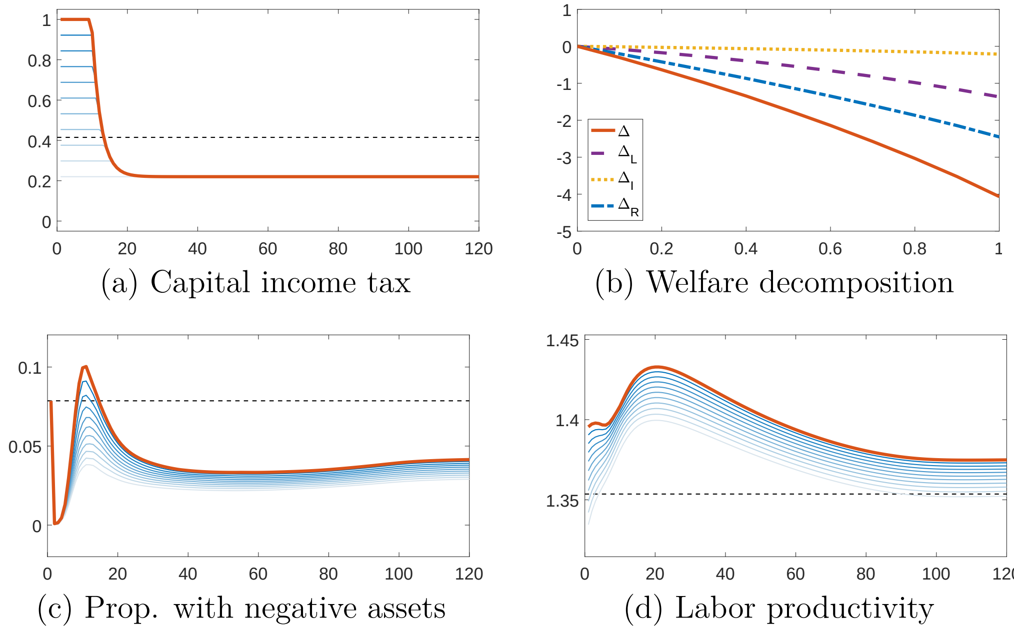
In Appendix G.1 we consider different levels of inequality aversion, but here we present results only for the extreme case in which the planner cares only about efficiency, namely $\hat{\sigma} = 0$.[30] Figure 11 presents the results in comparison with the benchmark results. Relative to the initial stationary equilibrium, the welfare gains associated with the policy are equivalent to a permanent 1.8 percent increase in consumption, 0.8 percent from reduction in distortions and 1.0 percent from extra insurance. Even though the planner does not take this into consideration, the policy implies a redistributive gain of about 1.1 percent.[31]

Relative to the benchmark experiment, capital labor income taxes are lower throughout the transition. Higher income taxes are beneficial both for insurance and redistributive motives, so it makes sense that removing one of these motives from consideration leads to lower levels of optimal income taxes.

Redistribution leads to efficiency gains. It is not at all obvious why it is optimal for maximizing efficiency to tax capital income at 100 percent for the first eight years. In a representative-agent setup without lump-sum taxes, the reason for front-loading capital

[30]The experiment of considering a planner that ignores redistributive concerns is similar to the experiment in Chari et al. (2018) restricting policies from reducing the value of initial wealth in utility terms, which effectively removes the planner's *possibility* to provide redistribution.

[31]Appendix K.2 contains the figures for aggregates associated with this experiment.



(a) Capital income tax (b) Welfare decomposition

(c) Prop. with negative assets (d) Labor productivity

Figure 12: Reducing Initial Capital Income Taxes

Note (a,c,d) Black dashed line: initial stationary equilibrium; Red solid and blue curves: path that maximizes efficiency and variations upon it; (b) the x-axis represents the homotopy parameter between the initial optimal path at $x = 0$ and a flat path at $x = 1$, y-axis shows change in the welfare gains in percent points.

income taxes is that the earlier the taxes are imposed, the less saving decisions are distorted. Here, the planner could reduce lump-sum transfers in every period, which would be distortive only to the extent that it brings households closer to their borrowing constraints. In Figure 12 we entertain exactly this experiment: we reduce the level of initial capital income taxes and decrease lump-sum in every period by the same amount to balance the budget.

First, notice, from Figure 12b that this hardly affects the insurance effect, although it does lead to a significant reduction in the level effect. This can be puzzling at first since it follows from a *reduction* in distortive taxes. Moreover, this variation actually reduces the proportion of households with negative assets (since capital income taxes subsidize negative asset holdings), so it is hard to argue the welfare losses are coming from forcing households toward their borrowing constraints. The key to make sense of these results is the increase in labor productivity, which follows from the redistribution achieved by the high initial capital income taxes. As explained above, redistribution generates wealth effects on labor supply that lead to a more efficient allocation of hours in the economy, with higher productivity households working relatively more—see Figure 12d. This effect is strong enough that it outweights the distortions associated with the high initial capital taxes.[32]

[32]This effect is not present in the earlier version of this paper, Dyrda and Pedroni (2016), since there we

Capital levy. An alternative way to investigate how much of the optimal policy has to do with redistribution is to consider an economy without initial inequality. In Appendix H, we present results for an experiment in which we remove the upper bound on capital income taxes. We show that, as a result, the planner completely expropriates the initial asset position of all households, removing all wealth inequality.[33] What is surprising, however, is that this actually leads to higher capital income taxes in future periods as well. This happens for three reasons: (1) in the short run, savings decisions are inelastic as households try to rebuild their buffer stocks of assets; (2) the large amount of assets acquired by the government crowds in capital, further mitigating distortions to capital accumulation; and (3) capital income taxes are still beneficial to provide redistribution (mostly in the short run) and insurance (mostly in the long run). Importantly, even though capital income taxes are overall higher relative to the benchmark, the equilibrium capital stock is still higher throughout the transition.

7 Effects of Transition and of Time-Varying Policies

In this section we quantify the importance of transitional effects and of the time variation of policy instruments. We first compute the optimal fiscal policy ignoring transitional welfare effects altogether. We show this leads to results that are substantially different from the benchmark. Then we solve for the optimal policy accounting for transitional welfare effects but subject to instruments remaining constant. The welfare loss associated with the constant-instrument constraint are still significant. The results are summarized in Table 5.

Table 5: Final Stationary Equilibrium: Transitional Effects

	τ^k	τ^h	T/Y	B/Y	K/Y	Δ	Δ_L	Δ_I	Δ_R
Initial equilibrium	41.5	22.5	11.4	61.5	2.49	–	–	–	–
Stat. equil.	–	36.4	18.8	−265.1	3.53	14.8	8.1	0.7	5.5
Stat. equil. no debt	−7.2	27.1	9.1	61.5	2.85	1.2	2.8	0.0	−1.5
Constant policy	67.5	27.9	19.7	53.9	2.02	1.7	−0.7	0.8	1.6
Benchmark	26.7	39.1	15.1	154.3	2.48	3.5	0.2	1.2	2.1

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

assume a utility function without wealth effects on labor supply.

[33]The expropriation of assets is combined with substantial lump-sum transfers in period 0, so that different savings in period 0 already bring the wealth Gini back to 0.25 by period 1.

7.1 Maximizing steady state welfare

Suppose the planner can choose stationary levels of all four fiscal policy instruments to maximize steady-state welfare. In particular, the planner can choose any level of government debt without incurring the transitional costs associated with it. It then chooses a debt-to-output ratio of -265 percent. At this level the amount of capital that is crowded in is close to the golden rule level, which implies zero interest rates (net of depreciation). Since capital income is zero, capital income taxes are not relevant which is why we do not display that number in Table 5. The average welfare gain associated with this policy is 14.8 percent. These are large welfare gains precisely because they ignore transitional effects, as if the economy has jumped immediately to a new steady state with a new distribution of assets, a much higher capital stock, and in which the government has a large amount of assets instead of debt.[34]

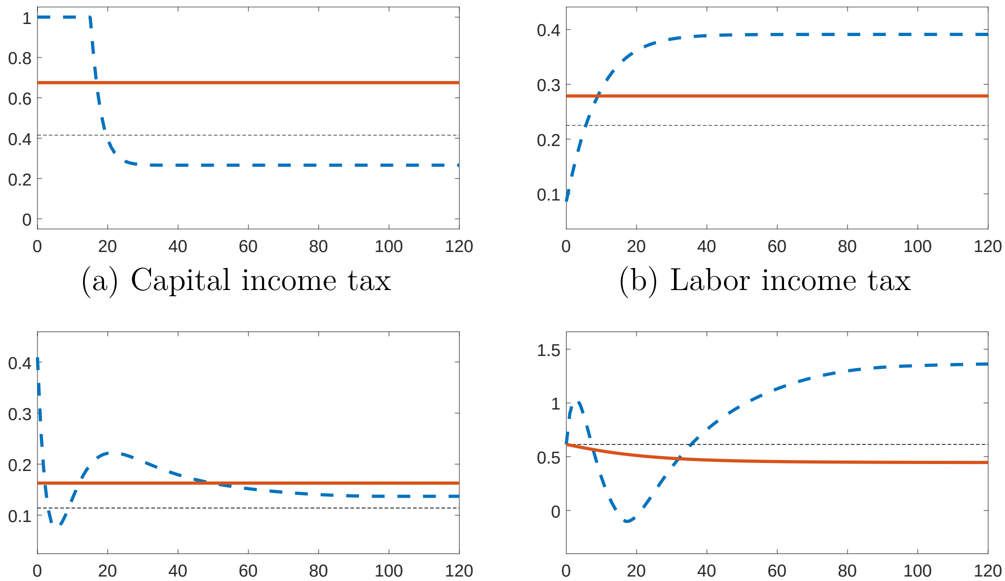
An alternative experiment, which is closer to the one studied by Conesa et al. (2009), is to restrict the level of debt-to-output to remain at its initial level and choose only the other fiscal instruments. With this constraint, we find it is still optimal for the planner to focus on the level effect. Though the golden rule level of capital is not achieved, a negative capital income tax of -7.2 leads the capital level in that direction. The planner also sets relatively low labor income tax and transfer levels which are detrimental to insurance and redistribution, but reduce distortions. Ignoring transitional effects, the policy leads to an average welfare gain of 1.2 percent. However, accounting for its transitional effects the policy would actually lead to a welfare *loss* equivalent to an 3.5 percent permanent reduction in consumption.

7.2 Transition with constant policy

If the planner does take transitional effects into account but cannot change taxes over time, the optimal fiscal instruments are set to a weighted average of the time-varying instruments from our benchmark results—see Figure 13.[35] More weight is put on the short-run levels since those are more relevant for welfare. The long-run levels of the fiscal instruments, however, are significantly different. Long-run capital income taxes and debt-to-output are 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

[34] In Appendix J.1 we compare these results with the ones in Aiyagari and McGrattan (1998) in detail. The main reason for the differences in the results is that the calibration in that paper leads to significantly lower levels of inequality.

[35] Figures with the corresponding aggregates are presented in Appendix K.3.



(a) Capital income tax (b) Labor income tax

(c) Lump-sum/Initial output (d) Debt/Initial output

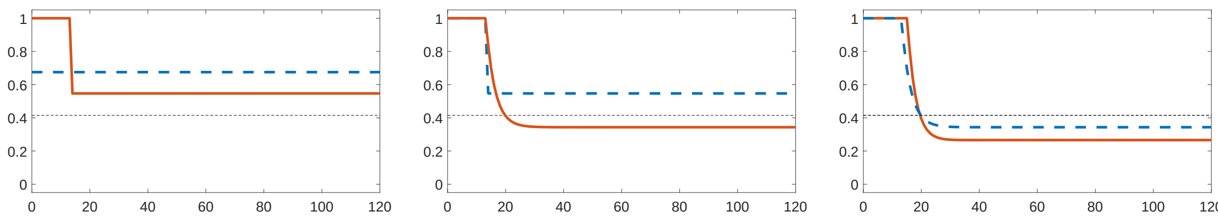
Figure 13: Optimal Fiscal Policy: Constant Policy

Notes: Black dashed line: initial stationary equilibrium; Red solid line or curve: path that maximizes efficiency optimal transition; Blue dashed curve: benchmark results.

particular, as we noticed above in Section 5.5, whereas the modified golden rule holds for the benchmark policy, it does not hold under this restriction—see Figure 10. Finally, this policy leads to welfare gains that are less than half those of the optimal dynamic policy.

7.3 Flexibility of time paths

To illustrate how increasing the flexibility of the parametrization of the paths of fiscal instruments affects the results we present, in Figure 14, some stages of this process for the path of capital income taxes. We are only showing the path for capital income taxes, but at each stage all fiscal instruments are allowed to follow more flexible paths and reoptimized. Figure 14a shows what happens when allow capital income taxes to be front-loaded: this minimal amount of flexibility increases welfare gains from 1.65 percent to 2.79. In Figure 14b, we show what happens when capital income taxes can follow the simplest form of equation (3.2), with only polynomials of degree zero, this involves choosing 8 parameters and improves welfare gains to 3.40 percent. Finally, Figure 14c shows what happens when we move from the 8-parameter solution to our benchmark 17-parameter solution, which brings welfare gains to 3.52 percent. At each step in which we add more flexibility, welfare increases by less. Nevertheless, some of the fiscal instruments still change in meaningful ways, as can be seen from the still significant reduction in long-run capital income taxes in Figure 14c. So, to determine optimal long-run policy accurately we keep adding flexibility until both welfare *and* policy are no longer affected. In Appendix G.3, we document all the intermediate steps of our implementation of this procedure with the corresponding figures and welfare gains.



(a) Front-Loading (b) More Flexibility (8 Par.) (c) Benchmark (17 Par.)

Figure 14: Adding Flexibility to Paths: Capital Income Taxes

Notes: Black dashed line: initial stationary equilibrium; Blue dashed curve in (a): optimal constant taxes; Red solid curve in (a) and blue dashed curve in (b): optimal transition allowing front-loading of capital income taxes; Red solid curve in (b) and blue dashed curve in (c): optimal transition with 8 parameters $(\alpha_0^k, \beta_0^k, \lambda^k, \alpha_0^h, \beta_0^h, \lambda^h, \beta_0^T, \lambda^T)$; Red solid curve in (c): benchmark optimal transition with 17 parameters—using $m_{kF} = m_{nF} = 0$, $m_{k0} = m_{n0} = m_{nF} = m_{TF} = 2$, and $m_{T0} = 4$ in equation (3.2).

8 Complete Market Economies

To understand how market incompleteness and different sources of inequality affect the optimal policy, we provide a build-up to our benchmark result. We start from a representative agent economy, without any heterogeneity whatsoever. Then, we introduce, labor-income and wealth inequality, in turn. Introducing uninsurable 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.

Importantly, for the complete market economies we can characterize the optimal policy analytically. We can also compute the optimal policy using this characterization and with the parameterized paths we used to obtain our benchmark results. The comparison between the two gives an idea of how well our numerical method approximates the actual optimal path. Notice that, in this complete-markets environment (without ad hoc borrowing constraints) the Ricardian equivalence holds, so the optimal paths for lump-sum taxes and debt are indeterminate, which is why we do not discuss or plot them.

The complete market economy is simply the SIM economy with the Markov transition matrix, Γ, set to the identity matrix and borrowing constraints replaced by no-Ponzi conditions. 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. Since the wealth distribution is indeterminate in the steady state of this economy, we can set the initial distribution to be the same as in our benchmark economy. We recalibrate the discount factor, β, to keep the same capital-to-output ratio.

Consider the same Ramsey problem as in Definition 3. With complete markets we can show that:

Proposition 6 *There exist a finite integer t^* and a constant Θ such that the optimal tax system is given by $\tau_t^k = 1$ for $0 \leq t < t^*$; while for $t \geq t^*$ τ_t^k follows*

$$\frac{1+(1-\tau_{t+1}^k)r_{t+1}}{1+r_{t+1}} = \frac{1-N_t}{1-N_{t+1}} \frac{1-\tau_{t+1}^h}{1-\tau_t^h} \frac{\tau_t^h+\tau^c}{\tau_{t+1}^h+\tau^c}; \tag{8.1}$$

for $0 \leq t \leq t^*$, τ_t^h evolves according to

$$\frac{1+(1-\tau_{t+1}^k)r_{t+1}}{1+r_{t+1}} = \frac{\Theta+\sigma\left(1-N_{t+1}\right)^{-1}}{\Theta+\sigma\left(1-N_t\right)^{-1}} \frac{1-\tau_{t+1}^h}{1-\tau_t^h} \frac{1+\tau^c+\alpha\left(\sigma-1\right)\left(\tau^c+\tau_t^h\right)}{1+\tau^c+\alpha\left(\sigma-1\right)\left(\tau^c+\tau_{t+1}^h\right)}; \tag{8.2}$$

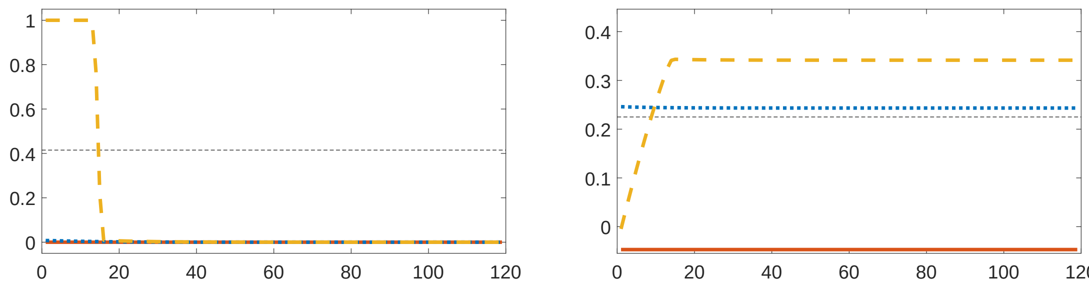
and for all $t > t^*$, τ_t^h is determined by

$$\tau_t^h\left(N_t\right) = \frac{\left(1+\tau^c\right)}{\left(1-N_t\right)\Theta+\alpha+\sigma\left(1-\alpha\right)} - \tau^c. \tag{8.3}$$

In Appendix F, we apply the method introduced by Werning (2007) to prove this proposition, and analogous ones for versions of this economy without labor–income and/or wealth inequality.[36] In particular, we also show that the magnitudes of t^* and Θ are related to the levels of wealth and labor–income inequality, respectively. Figure 15 illustrates the numerical results obtained using this proposition.

Representative agent. To avoid a trivial solution, Ramsey problems in a representative-agent economy usually do not allow lump-sum taxation. We do, so the solution in this case is indeed very simple. It is optimal to obtain all revenue via lump-sum taxes and set capital and labor income taxes so as not to distort any of the agent's decisions. This amounts to setting $\tau_t^k = 0$ and $\tau_t^h = -\tau^c$ for all $t \geq 0$. 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 the negative of the consumption tax ensures labor supply decisions are not distorted either.

[36] Werning (2007) allows complete expropriation of initial capital holdings. For comparability with our benchmark results, we impose an upper bound on capital income taxes and introduce an exogenous consumption tax.



(a) Capital income tax (b) Labor income tax

Figure 15: Optimal Taxes: Complete Market Economies

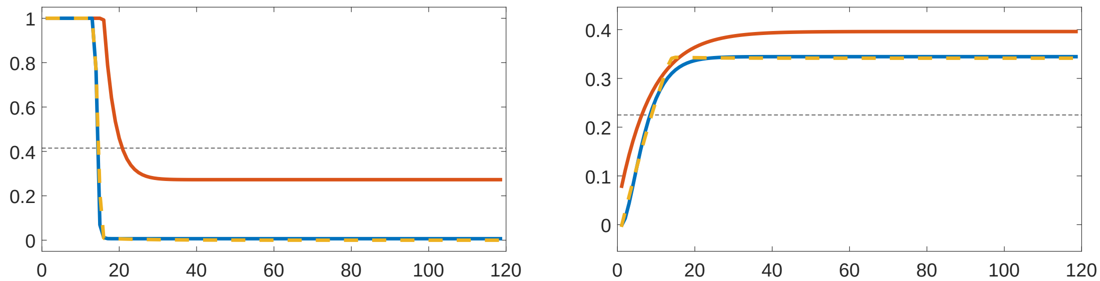
Notes: Black dashed line: initial taxes; Red solid curve: optimal taxes for representative economy; Dotted blue line: optimal taxes with only labor-income inequality; Yellow dashed curve: optimal taxes with labor-income and wealth inequality.

Labor-income inequality. When labor income is unequal, there is a redistributive reason to tax it. In Figure 15, we see that, in this case, it is optimal to have labor income taxes be virtually constant over time and capital income taxes virtually equal to zero in every period.

Wealth inequality. When there is wealth inequality there is a redistributive reason to tax asset income. With complete markets, however, capital income taxes are fully front-loaded, hitting the upper bound for t^* periods before converging to zero.[37] While capital income taxes are at the upper bound, labor income taxes are increasing. This leads to a decreasing (or less increasing) path for labor supply, which mitigates distortions to the households' intertemporal decisions: it leads to a smoother path for period utility as leisure increases while consumption decreases.

Uninsurable risk. Figure 16 contains the numerical results obtained using the same solution method used for the benchmark results together with the ones obtained using the proposition. This shows that, at least for this economy, the parameterized paths are able to approximated the actual solution relatively well (average welfare gains are similar as well: 2.253 percent using the proposition versus 2.246 percent using the parameterized paths). The figure also shows, for comparison, the results from the benchmark SIM model. The only important qualitative difference is the fact that for the SIM model capital income taxes are positive in the long run.

[37] Straub and Werning (2020) 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 Appendix F.8 we include a more detailed discussion of this issue.



(a) Capital income tax (b) Labor income tax

Figure 16: Optimal Taxes: Complete Market Economies

Notes: Black dashed line: initial taxes; Red solid curve: optimal taxes from Benchmark SIM model; Solid blue curve: optimal taxes calculated using the same parameterized paths used in the Benchmark experiment; Yellow dashed curve: optimal taxes calculated using Proposition 6.

9 Sensitivity Analysis and Robustness

In Appendix G we present the following robustness experiments: First, we show that higher degrees of inequality aversion for the planner are associated with higher taxes overall. However, particularly for values of inequality aversion above the benchmark utilitarian level, further increases have surprisingly small effects. Second, we show that changes in the IES have large effects specially on the path of optimal capital income taxes, because a different IES leads to a different relative risk aversion for households and a different degree of planner inequality aversion. The combined effect of all these changes can be large. Finally, we show that increases in the Frisch elasticity unsurprisingly reduce labor income taxes though by relatively small amounts.

In Appendices I and J, we present results for four alternative calibrations: (1) an economy that disciplines the labor income process without using any distributional moment, a common calibration strategy in the literature; (2) the calibration from Aiyagari and McGrattan (1998); (3) a calibration that introduces return-risk; and (4) the calibration from Acikgoz et al. (2018). There are two main takeaways from these experiments: the qualitative features of the Ramsey policy in the SIM model that we highlight in the paper—high short-run capital income taxes combined with increasing labor income taxes, and the front-loading of lump-sum transfers—are robust to substantial changes to the calibration; the quantitative results are sensitive to the calibration, which justifies the extensive effort we put into all the details of it.

10 Concluding Comments

In this paper, we quantitatively characterize the solution to the Ramsey problem in the standard incomplete markets model. We find that it is optimal to use distortive income taxes since they provide redistribution and insurance when rebated via lump-sum transfers—a utilitarian planner would expand the US social welfare system significantly, increasing overall transfers by roughly 50 percent. We quantify the associated welfare effects with a decomposition that accommodates transitional effects. We show that high initial capital income taxes are an effective way to provide redistribution, which also leads to a considerably more efficient allocation of labor via wealth effects on labor supply. Increasing labor income taxes over time mitigates the intertemporal distortions associated with high capital income taxes. Front-loading lump-sum transfers allows households to move away from their borrowing constraints making the government debt level less consequential for welfare, specially in the long run.

Finally, this paper abstracts from several important aspects that could be relevant for fiscal policy. For instance, in the model studied above, a household'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 the effects of international financial markets; capital income taxes as high as the ones we find optimal in this paper are unlikely to survive if households are able to move their assets overseas. We also abstract from life-cycle issues, and maintain 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.

References

Acikgoz, O. (2015): "Transitional Dynamics and Long-run Optimal Taxation Under Incomplete Markets," MPRA Paper 50160, University Library of Munich, Germany.

Acikgoz, O., M. Hagedorn, H. Holter, and Y. Wang (2018): "The Optimum Quantity of Capital and Debt," Tech. rep.

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

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

Aiyagari, S. R. and E. R. McGrattan (1998): "The optimum quantity of debt," *Journal of Monetary Economics*, 42, 447–469.

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

Bakis, O., B. Kaymak, and M. Poschke (2015): "Transitional Dynamics and the Optimal Progressivity of Income Redistribution," *Review of Economic Dynamics*, 18, 679–693.

Barsky, R. B., N. G. Mankiw, and S. P. Zeldes (1986): "Ricardian Consumers with Keynesian Propensities," *American Economic Review*, 76, 676–691.

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

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

Bewley, T. (1986): "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," in *Contributions to Mathematical Economics in Honor of Gerard Debreu*, ed. by W. Hildenbrand and A. Mas-Collel, Amsterdam: North-Holland.

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2017): "Public debt in economies with heterogeneous agents," *Journal of Monetary Economics*, 91, 39–51.

Boar, C. and V. Midrigan (2020): "Efficient Redistribution," Working Paper 27622, National Bureau of Economic Research.

Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818–857.

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

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

Chari, V. V., J. P. Nicolini, and P. Teles (2018): "Optimal Capital Taxation Revisited," Working Papers 752, Federal Reserve Bank of Minneapolis.

Chen, Y., C. C. Yang, and Y. Chien (2020): "Implementing the Modified Golden Rule?

Optimal Ramsey Capital Taxation with Incomplete Markets Revisited," Working Papers 2017-003, Federal Reserve Bank of St. Louis.

Conesa, J. C. and C. Garriga (2008): "Optimal Fiscal Policy in the Design of Social Security Reforms," *International Economic Review*, 49, 291–318.

Conesa, J. C., S. Kitao, and D. Krueger (2009): "Taxing Capital? Not a Bad Idea after All!" *American Economic Review*, 99, 25–48.

Conesa, J. C. and D. Krueger (1999): "Social Security Reform with Heterogeneous Agents," *Review of Economic Dynamics*, 2, 757–795.

Dávila, J., J. H. Hong, P. Krusell, and J.-V. Ríos-Rull (2012): "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks," *Econometrica*, 80, 2431–2467.

Domeij, D. and J. Heathcote (2004): "On The Distributional Effects Of Reducing Capital Taxes," *International Economic Review*, 45, 523–554.

Dyrda, S. and M. Pedroni (2016): "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks," Working Papers 1245, Society for Economic Dynamics.

Erosa, A., L. Fuster, and G. Kambourov (2016): "Towards a Micro-Founded Theory of Aggregate Labour Supply," *Review of Economic Studies*, 83, 1001–1039.

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

Gottardi, P., A. Kajii, and T. Nakajima (2015): "Optimal Taxation and Debt with Uninsurable Risks to Human Capital Accumulation," *American Economic Review*, 105, 3443–70.

——— (2016): "Constrained Inefficiency and Optimal Taxation with Uninsurable Risks," *Journal of Public Economic Theory*, 18, 1–28.

Greulich, K., S. Laczó, and A. Marcet (2019): "Pareto-Improving Optimal Capital and Labor Taxes," Tech. rep.

Guvenen, F. (2011): "Macroeconomics with hetereogeneity : A practical guide," *Economic Quarterly*, 97, 255–326.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework," *The Quarterly Journal of Economics*, 132, 1693–1754.

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

——— (1997): "The one-sector growth model with idiosyncratic shocks: Steady states and dynamics," *Journal of Monetary Economics*, 39, 385–403.

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

Itskhoki, O. and B. Moll (2019): "Optimal Development Policies With Financial Frictions," *Econometrica*, 87, 139–173.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1997): "On the Optimal Taxation of Capital

Income," *Journal of Economic Theory*, 73, 93–117.

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

——— (2002): "The parametric path method: An alternative to Fair–Taylor and L–B–J for solving perfect foresight models," *Journal of Economic Dynamics and Control*, 26, 1557 – 1583, special issue in honour of David Kendrick.

Kan, A. R. and G. T. Timmer (1987a): "Stochastic Global Optimization Methods Part I: Clustering Methods," *Mathematical Programming*, 39, 27–56.

——— (1987b): "Stochastic Global Optimization Methods Part II: Multilevel Methods," *Mathematical Programming*, 39, 57–78.

Kindermann, F. and D. Krueger (2014): "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," PIER Working Paper Archive 14-036, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Krueger, D. and A. Ludwig (2016): "On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium," *Journal of Monetary Economics*, 77, 72–98.

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

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

Nakajima, M. and J.-V. Ríos-Rull (2019): "Credit, Bankruptcy, and Aggregate Fluctuations," Working Papers 19-48, Federal Reserve Bank of Philadelphia.

Nuño, G. and C. Thomas (2016): "Optimal monetary policy with heterogeneous agents (Updated September 2019)," Working Papers 1624, Banco de España; Working Papers Homepage.

Ponce, M., R. van Zon, S. Northrup, D. Gruner, J. Chen, F. Ertinaz, A. Fedoseev, L. Groer, F. Mao, B. C. Mundim, M. Nolta, J. Pinto, M. Saldarriaga, V. Slavnic, E. Spence, C.-H. Yu, and W. R. Peltier (2019): "Deploying a Top-100 Supercomputer for Large Parallel Workloads: The Niagara Supercomputer," in *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)*, New York, NY, USA: Association for Computing Machinery, PEARC '19.

Powell, M. (2009): "The BOBYQA algorithm for bound constrained optimization without derivatives," Tech. rep., Department of Applied Mathematics and Theoretical Physics, Cambridge University.

Pruitt, S. and N. Turner (2020): "Earnings Risk in the Household: Evidence from Millions of US Tax Returns," *American Economic Review: Insights*, 2, 237–54.

Ragot, X. and F. L. Grand (2020): "Optimal Fiscal Policy with Heterogeneous Agents and Aggregate Shocks," Tech. rep.

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

Rouwenhorst, G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," In Frontiers of Business Cycle Research, edited by Thomas F. Cooley. Princeton University Press.

Saez, E. and S. Stantcheva (2018): "A simpler theory of optimal capital taxation," *Journal of Public Economics*, 162, 120–142.

Straub, L. and I. Werning (2020): "Positive Long-Run Capital Taxation: Chamley–Judd Revisited," *American Economic Review*, 110, 86–119.

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

Trabandt, M. and H. Uhlig (2012): "How Do Laffer Curves Differ across Countries?" in *Fiscal Policy after the Financial Crisis*, National Bureau of Economic Research, Inc., NBER Chapters, 211–249.

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