Review of Economic Studies (2023) **90**, 744–780 https://doi.org/10.1093/restud/rdac031

Advance access publication 2 June 2022

# Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk

SEBASTIAN DYRDA
*University of Toronto*

and

MARCELO PEDRONI
*University of Amsterdam*

*First version received December* 2020*; Editorial decision January* 2022*; Accepted May* 2022 (*Eds.*)

We study optimal fiscal policy in a standard incomplete-markets model with uninsurable idiosyncratic income risk, where a Ramsey planner chooses time-varying paths of proportional capital and labour 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 labour income taxes over time mitigates intertemporal distortions from capital income taxes; (3) the optimal policy increases overall transfers, calibrated initially to the US welfare system, by roughly 50%; (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 labour via wealth effects on labour supply—lower productivity households can afford to work relatively less.

*Key words*: Optimal taxation; Heterogeneous agents; Incomplete markets

*JEL Codes*: E2, E6, H2, H3, D52

## 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 labour 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 labour income process of

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

*The editor in charge of this paper was Dirk Krueger.*

households. We then consider a Ramsey planner that finances an exogenous stream of government expenditures with proportional capital and labour income taxes, lump-sum transfers (or taxes), and government debt. We allow policy to be *time varying* and evaluate welfare over the *transition*. 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.

We find that a utilitarian planner would confiscate capital income for the initial 16 years, and still tax it at a positive rate of 27% in the long run, lower than the prevailing rates in the US of 42%. Labour income taxes increase over time in the 16 initial years reaching 39% in the long run, a significantly higher level than the prevailing rate of 23%. These changes in income taxes are used to finance an increase in lump-sum transfers of roughly 50% on average over time. At the same time, the ratio of government debt to GDP more than doubles to 154% in the long run. This policy leads to welfare gains equivalent to a permanent increase in consumption of 3.5%.

More generally, we provide new insights about the dynamics of the optimal policy in the SIM model. The initial confiscation of capital income, rebated via lump-sum transfers, is effective in providing redistribution, since the tax base is relatively unequal and inelastic. The resulting distortions to the intertemporal margin are mitigated by an increasing path of labour income taxes this period and, in a subtle way, by a non-monotonic path of lump-sum transfers. The achieved redistribution also activates a wealth effect on labour supply that leads to a more efficient allocation of labour, increasing the correlation between productivity and hours worked. The overall more generous tax-transfer system also provides insurance to the income risk faced by households. 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. The average welfare gains of 3.5% can be decomposed into: (1) 0.2% from a reduction in distortions to households' decisions, (2) 1.2% from insurance (the reduction of *ex post* risk), and (3) 2.1% from redistribution (the reduction of *ex ante* risk). 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 labour income are both unequally distributed between households and risky over time. Labour and capital income taxes distort households' savings and labour 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 (IES) 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. a planner that minimizes distortions—or maximizes the welfare of the average household—and minimizes risk faced by households given their initial conditions). 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 labour supply, to a relative increase in hours worked by the more productive. The end result is a substantial increase in average labour 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 should emphasize that the complementary between efficiency and redistribution hinges on the strength of wealth effects on labour supply and disappears when these are set to zero (as implied by GHH preferences), so we are careful to discipline these wealth effects well by matching at the same time the distributions of earnings, wealth, and hours worked.

We also show that the time variation of fiscal instruments is important. If they are restricted to being constant over time, the welfare gains are roughly *half* of the ones implied by the optimal policy, in large part because the movements over time allow the cross mitigation of distortions. Time variation is also crucial if one is interested in 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 labour productivity rationalizes distortive labour 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, labour 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 labour 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. Nevertheless, we find that the optimal time variation of lump-sum transfers and debt contribute only marginally to the overall welfare gains. Specifically, reoptimizing subject to the constraint that lump-sum transfers be constant over time, or that the debt-to-output must be fixed at its pre-reform level, leads to welfare losses of about 0.1 and 0.2%, respectively. There are three reasons for this: (1) departures from Ricardian equivalence are quantitatively relevant in proportion to how close households are to their borrowing constraints; (2) under the optimal policy only a minority of households are borrowing constrained; and (3) the general equilibrium price effects associated with changes in debt have counteracting effects on redistribution and insurance.

### 1.1. *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 (MGR) 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. Moreover, Acikgoz *et al.* (2018) show that long-run fiscal policy can be characterized independently of initial conditions and solve backwards for the optimal transition. We offer an alternative method of solving for the optimal policies in the SIM model, which does not require establishing the independence of the long-run policies from

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 IES.

transitional dynamics and which can be applied to any model in which one can compute transitions fast enough, even if first-order conditions are not tractable.[3]

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 analysing 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.

This article is also related to the emerging literature on universal basic income—Guner, Kaygusuz and Ventura (2021), Luduvice (2019) and Daruich and Fernández (2020). Our measurement of lump-sum transfers covers all sources of transfers provided by the federal government which imply a lower bound to income. The overall increase in lump-sum transfers suggested by the Ramsey policy could be implemented by the introduction of an universal basic income.

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 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 labour income process and compute optimal government debt not only in the long run but

3. We also extend the results 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. We discuss the relationship between our method and results and theirs in Section 5.6 and, in more detail, in Supplementary Appendix M.

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 of the SIM model.

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 the price effects of changes in government debt. This mechanism plays a role in some of our results.

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), but 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 consider a general-equilibrium two-period economy to explore how exogenous changes to risk and inequality affect the optimal tax system. We show that the presence of uninsurable labour-productivity risk creates a reason to use distortive labour income taxes even if the planner is able to obtain all necessary revenue using the undistortive lump-sum instrument. Similarly, we show that more inequality leads to higher optimal levels of capital income taxes. These takeaways are useful to interpret the results in the more complicated quantitative model that follows.

### 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 labour, $\gamma$ controls the consumption share, and $\sigma$ controls the preference for risk and over-time smoothness. Also, suppose that households discount the future by a factor of $\beta$.

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

In Period 2, output is produced using capital, $K$, and labour, $N$, and a constant-returns-to-scale neoclassical production function $F(K,N)$ which includes undepreciated capital. The government

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 Supplementary Appendix F.8.

needs to finance an expenditure of $G$. It has three instruments available: labour income taxes, $\tau^h$, capital taxes, $\tau_R^k$,[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* $N=\pi_Le_Lh_L+\pi_He_Hh_H$;
3. *and,* $\tau^hwN+\tau_R^kRK=G+T$.

The Ramsey problem is to choose $\tau^h$, $\tau_R^k$, 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 non-distortive 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*

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

*Further,* $\Omega=0$ *if* $e_L=e_H$*, and for an increase in risk via a mean-preserving spread* $\varepsilon$*, such that productivities become* $(e_L-\varepsilon/\pi_L, e_H-\varepsilon/\pi_H)$*, we have that* $\partial\Omega(\varepsilon)/\partial\varepsilon>0$.

The proofs of the results in this section can be found in Supplementary Appendix B.[10] Notice that $\Omega$, which is an endogenous object, 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 follows from the formula for $\tau^h$ that labour 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.[11] 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

8. Below, we denote capital *income* taxes by $\tau^k$, but here it is more convenient to use $\tau_R^k$.

9. In a similar two-period environment, Gottardi, Kajii and Nakajima (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. Supplementary 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. This mechanism is reminiscent of Barsky, Mankiw and Zeldes (1986).

labour taxes via wealth effects: taxing capital reduces wealth in Period 2 which increases labour supply.[12]

### 2.2. *The effect of inequality*

Consider the environment described above 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: $\omega_L$ for a proportion $p_L$ of households, and $\omega_H > \omega_L$ for the rest. Let $\bar{\omega}$ denote the 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 $\sigma = 1$,[13] then the utilitarian optimal tax system is such that*

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

*where*

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

*Further, $\Lambda = 0$ if $\omega_L = \omega_H$, and for an increase in inequality via a mean-preserving spread $\varepsilon$, such that the initial endowments become $(\omega_L - \varepsilon/p_L, \omega_H - \varepsilon/p_H)$, we have that $\partial\Lambda(\varepsilon)/\partial\varepsilon > 0$.*

Here, $\Lambda$, which is also endogenous, 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, rebated via lump-sum transfers, directly reduces the proportion of household income that depends on unequal asset income achieving the desired redistribution.

Optimal labour income taxes are set to zero. To see why, consider increasing labour taxation and rebating the extra revenue via a lump sum. Since asset-poorer households have a higher proportion of their income coming from labour, 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. Labour income taxes affect not only the level of risk through the mechanism described above but also labour 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.

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

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

Therefore, capital income taxation affects both *ex ante* and *ex post* risk faced by households. Nevertheless, these results are useful for understanding some 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 $\mathbb{E}_0\left[\sum_t \beta^t u(c_t, h_t)\right]$, where $c_t$ and $h_t$ denote consumption and hours worked in period $t$. The household's labour productivity, denoted by $e \in E$ with $E \equiv \{e_1, \ldots, e_L\}$, follows a Markov process governed by the transition matrix $\Gamma$. 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 then indexed by the pair $(e,a) \in S$. Given a sequence of prices $\{r_t, w_t\}_{t=0}^{\infty}$, labour 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 $c_t(a,e)$, $h_t(a,e)$, and $a_{t+1}(a,e)$ 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 both the value and the 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, $\tau^c$, is a parameter.[14] Let $\{\lambda_t\}_{t=0}^{\infty}$ be a sequence of probability measures over the Borel sets $\mathcal{S}$ of $S$ with $\lambda_0$ given. Since the path for taxes is known, prices and$\{\lambda_t\}_{t=0}^{\infty}$ follow deterministic paths. As a result, we do not need to keep track of the distribution as an additional state; time is a sufficient statistic.

Competitive firms own a constant-returns-to-scale technology $f(\cdot)$ that uses capital, $K_t$, and efficient units of labour, $N_t$, to produce output each period: $f(\cdot)$ denotes output net of depreciation, while $\delta$ is the depreciation rate. A representative firm exists that solves the usual static problem. The government needs to finance an exogenous constant stream of expenditure, $G$, and lump-sum transfers with taxes on consumption, labour income, and capital income. The government can also issue debt, $\left\{B_{t+1}\right\}_{t=0}^{\infty}$, 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.

14. It is not without loss of generality that we do not allow the planner to choose $\tau^c$. There are two reasons for this choice. The first is practical: we are already using the limit of the computational power available to us, and allowing for one more choice variable would increase it substantially. Second, in the US, capital and labour 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 $\tau^c$ as a parameter for calibration purposes.

**Definition 2.** *Given $K_0$, $B_0$, an initial distribution $\lambda_0$, and a policy $\pi \equiv \{\tau_t^k, \tau_t^h, T_t\}_{t=0}^{\infty}$, a* ***competitive equilibrium*** *is a sequence of value functions $\{v_t\}_{t=0}^{\infty}$, an allocation $X \equiv \{c_t, h_t, a_{t+1}, K_{t+1}, N_t, B_{t+1}\}_{t=0}^{\infty}$, 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 $\pi$, $c_t(a,e)$, $h_t(a,e)$, 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*

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

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

4. *The government budget constraint, (3.1), holds and debt is bounded;*[15]
5. *Markets clear,*

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

### 3.1. *The Ramsey problem*

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

**Definition 3.** *Given $K_0$, $B_0$, and $\lambda_0$, for every policy $\pi$,* ***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 $\Pi$ 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 $\lambda_0$. The planner's objective is, thus, given by

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

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 finite-dimensional one, we assume the existence

15. We do not impose any exogenous upper bound on the path of government debt. By "debt is bounded" we mean that there exists $M$ such that $|B_t| < M$ for every $t \geq 1$, but we do not specify any $M$.

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.[16] To decrease the dimensionality of the problem further, we build on Judd (2002) and parameterize the time paths of fiscal instruments as follows:

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

where $x_t$ can be any of the fiscal instruments $\tau_t^k$, $\tau_t^h$, 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; $\left\{\alpha_i^x\right\}_{i=0}^{m_{x0}}$ and $\left\{\beta_j^x\right\}_{j=0}^{m_{xF}}$ are weights on the consecutive elements of the family; $\lambda^x$ 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, 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 initial polynomial families of degree two for labour and capital income taxes ($m_{\tau^k 0} = m_{\tau^h 0} = 2$), and four for lump-sum transfers ($m_{T0} = 4$), and final polynomial families of degree zero for labour and capital income tax ($m_{\tau^k F} = m_{\tau^h F} = 0$) and two for lump-sum transfers ($m_{TF} = 2$).[17] We set the terminal period at which taxes become constant to be $t_F = 100$,[18] and an upper bound on the capital income taxes of $\bar{\tau}^k = 1$, following the Ramsey literature.[19] Given these choices, we end up with the following 17 parameters:

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

which determine the time paths of fiscal instruments.

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).

16. By stationary equilibrium, we mean that all objects in Definition 2 become time-invariant. We should note that while the assumption of the existence of a Ramsey steady state is common in the literature it may not be innocuous as exemplified by Straub and Werning (2020). The specific issue highlighted by Straub and Werning (2020), however, is not a problem in our setup as a result of lump-sum transfers being available to the planner, see Supplementary Appendix F.8 for more details.

17. In Supplementary Appendix G.3, we discuss how the optimal policy changes as we gradually increase the number of choice variables.

18. 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 Supplementary 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. We also recomputed the optimal policy increasing the length of the transition from 250 to 500 and obtained essentially identical results.

19. In Supplementary Appendix O.6, we show how the policy is affected for different choices for $\bar{\tau}^k$, whereas in Supplementary Appendix I, we consider the case without any upper bound.

A detailed description is contained in Supplementary 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 $\pi_A$. 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 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.[20]

## 4. CALIBRATION

A period in the model is considered to be 1 year. 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: (1) time series of macroeconomic data from 1995 to 2007, (2) cross-sectional, distributional moments on hours worked, wealth, and earnings, and (3) panel data on the dynamics of labour 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 44 targets to discipline them, so the system is overidentified. Parameter values, targeted statistics, and their model counterparts are presented in Tables 1 and 2. Supplementary Appendix A contains a detailed description of how we calculated the targets from the data.

### 4.1. *Households vs. 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, $\beta$, is chosen to match a capital-output ratio of 2.5.[21] The two parameters in the balanced-growth-path utility function (2.1), $\gamma$ and $\sigma$ are disciplined with two targets: (1) an IES of 0.65, which sits between the numbers used in the related literature of 0.5 in Conesa *et al.* (2009) and Dávila *et al.* (2012), 0.8 in Straub and Werning (2020) and 0.86 in Aiyagari and McGrattan (1998), and implies a relative risk aversion of 1.55;[22] and (2) the average hours worked of

20. The baseline experiment was conducted using 1200 cores on the Niagara supercomputer at the University of Toronto, see Ponce, van Zon, Northrup, Gruner, Chen, Ertinaz, Fedoseev, Groer, Mao, Mundim, Nolta, Pinto, Saldarriaga, Slavnic, Spence, Yu and Peltier (2019) and Supplementary Appendix D.3 for details about the cluster.

21. Capital is defined as non-residential and residential private fixed assets and purchases of consumer durables. For more details, see Supplementary Appendix A.1.

22. Relative to the more conventional IES of 0.5, our choice of 0.65 is also an attempt to absorb, to some extent, new relevant empirical findings. Recent empirical evidence has generally pointed to higher IES levels (e.g. Bansal and Yaron, 2004; Hansen, Heaton, Lee and Roussanov, 2007; Barro, 2009; Bansal, Kiku and Yaron, 2012; Gruber, 2013) and lower CRRA levels (see Chetty, 2006). In Supplementary Appendix G.2, we show that we can achieve otherwise very similar calibration results with an IES of 0.5 or 0.8, and, in Section 9, we conduct a sensitivity analysis with respect to this choice.

TABLE 1
*Benchmark model parameters*

| **Description** | **Parameter** | **Value** | | |
|---|---|---|---|---|
| Preferences and technology | | | | |
| Consumption share | $\gamma$ | 0.510 | Implied IES $\left(\frac{1}{1-\gamma(1-\sigma)}\right)$: | 0.65 |
| Preference curvature | $\sigma$ | 2.069 | Implied Frisch ($\Psi$): | 0.49 |
| Discount factor | $\beta$ | 0.954 | | |
| Capital share | $\alpha$ | 0.378[a] | | |
| Depreciation rate | $\delta$ | 0.104 | | |
| Borrowing constraint | $\underline{a}$ | −0.078 | | |
| Fiscal policy | | | | |
| Capital income tax (%) | $\tau^k$ | 41.5[a] | | |
| Labour income tax (%) | $\tau^h$ | 22.5[a] | | |
| Consumption tax (%) | $\tau^c$ | 4.7[a] | | |
| Government expenditure | $G$ | 0.069 | | |
| Transfers | $T$ | 0.088 | | |
| Labour productivity process | | | | |
| Productivity process curvature | $\eta$ | 1.153 | | |
| Persistent shock | | | Transitory shock | |
| $\Gamma_P = \begin{bmatrix} 0.994 & 0.002 & 0.004 & 3\text{E}{-5} \\ 0.019 & 0.979 & 0.001 & 9\text{E}{-5} \\ 0.023 & 0.000 & 0.977 & 5\text{E}{-5} \\ 0.000 & 0.000 & 0.012 & 0.987 \end{bmatrix}$ | $e_P = \begin{bmatrix} 0.580 \\ 1.153 \\ 1.926 \\ 27.223 \end{bmatrix}$ | | $P_T = \begin{bmatrix} 0.263 \\ 0.003 \\ 0.556 \\ 0.001 \\ 0.001 \\ 0.176 \end{bmatrix}$ | $e_T = \begin{bmatrix} -0.574 \\ -0.232 \\ 0.114 \\ 0.133 \\ 0.817 \\ 1.245 \end{bmatrix}$ |

*Notes:* [a]Exogenously set parameters.

employed households in the Current Population Survey (CPS) between 1995 and 2007, which is equal to 0.32.

To discipline the extensive margin labour-supply decision we target the fraction of employed households in the economy. We follow Heathcote, Perri and Violante (2010) and consider a household to be employed if they work more than five hours per week, that is, if $h \geq \underline{h} \equiv 0.05 = 260/52,000$. Using data from the CPS, we calculate that 79% of households are employed—see Supplementary Appendix A.3 for more details. Since household-level Frisch elasticities depend on the household's labour supply, we measure the intensive-margin aggregate Frisch elasticity with the unweighted average of household-level Frisch elasticities for employed households, that is,

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

Our calibration implies a value for $\Psi$ of 0.49 which is close to the 0.54 reported by Chetty, Guren, Manoli and Weber (2011) in their survey of estimates of the Frisch elasticity.[23] We conduct sensitivity analysis with respect to our choice for the IES and this measure of Frisch elasticity in Section 9. The values of preference parameters, together with the implied elasticities

23. To check whether the extensive-margin elasticity of labour supply is also in line with the data, we consider the transitional dynamics following a temporary 1% 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).

TABLE 2
*Benchmark model economy: target statistics and model counterparts*

| | Target | Model |
|---|---|---|
| (1) Macroeconomic aggregates | | |
| Intertemporal elasticity of substitution | 0.65 | 0.65 |
| 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 |
| Fraction of employed (%) | 79.0 | 80.4 |
| 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 | 1.8 | 3.7 | 8.9 | 84.3 | 56.3 | 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 | 5.5 | 10.5 | 19.7 | 62.3 | 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 | 12.9 | 22.4 | 25.7 | 35.0 | 9.8 | 0.36 |

| | Target | Model |
|---|---|---|
| (3) Statistical properties of labour income | | |
| Variance of 1-year growth rate | 2.3 | 2.2 |
| Kelly skewness of 1-year growth rate | −0.1 | −0.1 |
| Moors kurtosis of 1-year growth rate | 2.7 | 2.3 |
| (4) Self-employed statistics | | |
| Share in population (%) | 12.5 | 12.7 |
| Share of wealth (%) | 45.8 | 38.9 |
| Share of earnings (%) | 28.7 | 30.5 |

are reported in the first three rows of Table 1, while the targets disciplining them are presented in the first three rows of Table 2.

The production function, net of depreciation, is given by $f(K,N)=K^{\alpha}N^{1-\alpha}-\delta K$. The depreciation rate, $\delta$, is set to match an investment-to-output ratio of 26%, and the capital share, $\alpha$, to its empirical counterpart of 0.38.[24] These choices imply an interest rate of 4.7%. 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%.

## 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 the initial capital income tax to

24. These numbers are computed in a consistent way with the capital-output ratio, and Supplementary Appendix A.1 describes their calculation in detail.

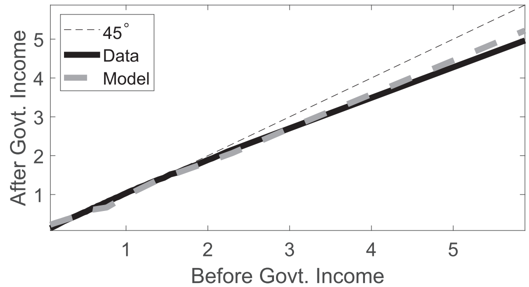


FIGURE 1
Income tax schedule

*Notes:* The data were 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.

41.5%, the labour income tax to 22.5%, and the consumption tax to 4.7%. We discipline the lump-sum transfer by targeting the average transfer-to-output ratio in the US from 1995 to 2007, which amounts to 11.4%.[25] We set the government debt-to-output ratio in the initial equilibrium to be 61.5%, averaging out federal debt over GDP in the data from 1995 to 2007. These choices of fiscal parameters are summarized in the rows labelled "Fiscal policy" in Table 1 and "Macroeconomic aggregates" in Table 2. The calibrated values implies a government-expenditure-to-output ratio of 8.9%, while the data counterpart (federal government expenditure) for the relevant period is approximately 6.9%. Further, we closely approximate the actual income tax schedule—see Figure 1.

### 4.4. *Labour productivity process*

The stochastic process for household labour productivity levels, $e$, is calibrated to match statistical properties of the labour income process as well as the cross-sectional distributions of hours worked, wealth, and earnings. The productivity levels have a persistent component $e_P$ with Markov matrix $\Gamma_P$, and a transitory component $e_T$ with probability vector $P_T$.[26] There are four persistent and six transitory productivity levels. We normalize the average productivity to one, so we are left with 26 free parameters associated with the labour income process.

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 Rouwenhorst (1995), and target a small set of moments usually only focusing on the labour-income process itself. The second approach, put forward by Castañeda, Díaz-Giménez and Ríos-Rull (2003), abstracts from labour 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 target moments of the labour income process itself, including higher moments such as skewness and kurtosis of their growth rates. This gives us the ability to match, at the same time, important measures of inequality and risk faced by households. The transition matrix governing the persistent

25. 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. For more details, see Supplementary Appendix A.1.

26. In the notation of the model, $\Gamma = \Gamma_P \otimes \text{diag}(P_T)$, 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.

TABLE 3
*Income sources of households by quintile of wealth*

| Quintile | Model | | | US data | | |
|---|---|---|---|---|---|---|
| | Labour | Asset | Transfer | Labour | Asset | Transfer |
| 1st | 80.2 | 0.2 | 19.8 | 83.6 | 0.4 | 16.1 |
| 2nd | 77.0 | 2.6 | 20.4 | 86.5 | 1.1 | 12.3 |
| 3rd | 74.4 | 5.3 | 20.3 | 85.6 | 1.9 | 12.5 |
| 4th | 74.8 | 9.4 | 15.7 | 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 |

*Notes*: This table summarizes the pre-tax total income decomposition. The data comes Table 6 in Díaz-Giménez *et al.* (2011) who summarize the 2007 SCF. We define total income using all categories but "Other." We split "Business" income into labour and asset income using the proportion of overall "Labour" to "Capital" income.

shocks, the probabilities associated with transitory shocks, and the corresponding productivity levels are reported in Table 1 under the "Labour productivity process" label.

**4.4.1. Inequality.** We target the share owned by every quintile, the Gini coefficient, and the share owned by the bottom and top 5% of the wealth, earnings, and hours distributions. For wealth and earnings, we use data from the SCF, and for hours we use the CPS. We report the performance of the model with respect to these targets in Table 2 under the label "Cross-sectional distributions." To account for the joint distribution of earnings and wealth, we also target the cross-sectional correlation between them.

**4.4.2. Risk.** Pruitt and Turner (2020) document statistical properties of the labour 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 labour income, which we target. We report them in Table 2 under the "Statistical properties of labour income" label. 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, the 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 labour-income moments.[27] The targeted moments for entrepreneurs, together with their model counterparts are reported in Table 2 under the label "Self-employed statistics."

### 4.5. *Model performance*

Table 3 presents income sources over quintiles of income. The composition of income, especially of consumption-poor households, plays an important role in determining the optimal fiscal policy. The fraction of uncertain labour 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. Figure 2 presents how well the model matches the targeted cross-sectional distributions of wealth, earnings, and hours. The last two panels of the figure show that the model also approximates well

27. A similar strategy has been employed by Kindermann and Krueger (2021) and Nakajima and Ríos-Rull (2019).

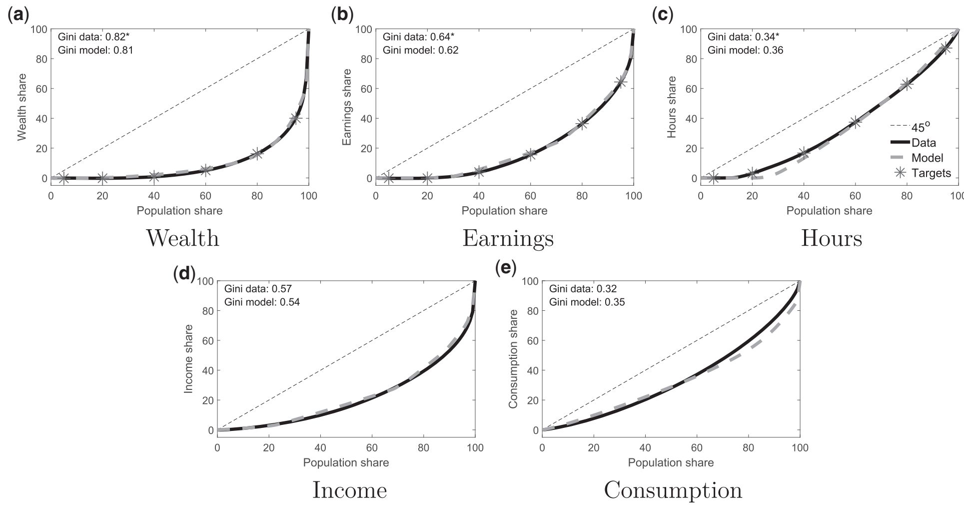


FIGURE 2
Fit to inequality data

the untargeted distributions of income and consumption. The earnings elasticity of the most productive households plays a role in some of the arguments we present below. So, we followed the procedure in Kindermann and Krueger (2021) to calculate this elasticity for the top 1%. The elasticity in the model implies that the peak of the Laffer curve lies at 78%, which is reasonably close to their targeted value of 73%—see Supplementary Appendix K for more details.

## 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% in the long run. The labour income tax drops on impact to 9% and then monotonically increases to 39% in the long run. Lump-sum transfers jump to 40% of output on impact, follow a U-shaped pattern in the short-run and, starting from a period 22, fall monotonically toward 15% of output in the long run. The government debt-to-output ratio 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% of output in the long run. This policy yields welfare gains equivalent to a 3.5% permanent increase in the consumption of all households.

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*

Figure 4 summarizes the main effects of the optimal policy on aggregates.[28] High capital income taxes in the initial periods lead to a reduction in the capital stock of about 10%. The substantial

28. Supplementary Appendix O.1 contains a more comprehensive list of figures.

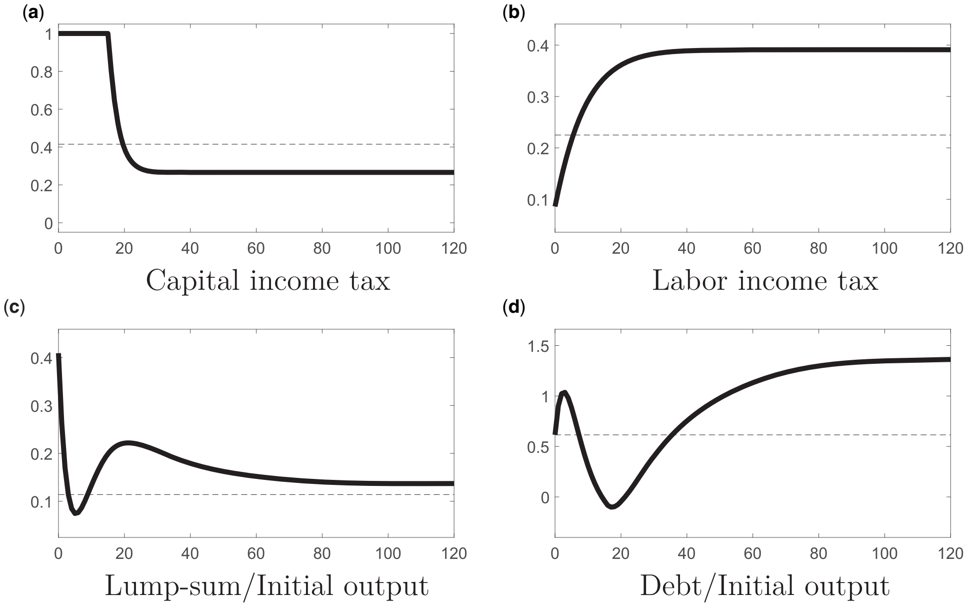


FIGURE 3

Optimal fiscal policy: benchmark

*Notes:* Thin dashed lines: initial stationary equilibrium; Thick solid curves: optimal transition.

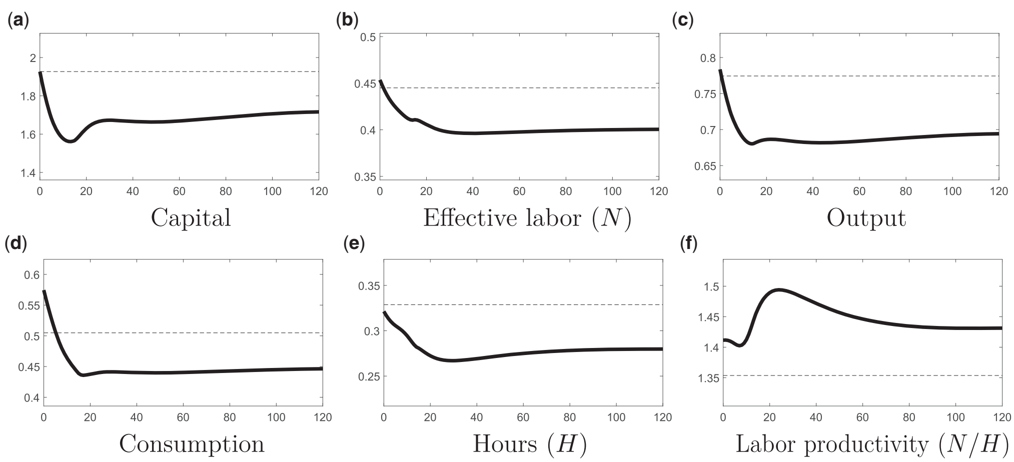


FIGURE 4

Optimal fiscal policy: aggregates

*Notes:* Thin dashed lines: initial stationary equilibrium; Thick solid curves: optimal transition.

fall in these taxes later on does not imply a recovery for three reasons: (1) government debt increases, which crowds out private capital, (2) labour decreases over time as a result of higher labour income taxes, which reduces the marginal 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% 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 labour.

Even with lower labour 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 labour supply reduce the labour supply of the more numerous lower-productivity households. The subsequent reduction in hours worked are due to increasing labour income taxes. In the long run, aggregate hours fall by 15% 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 the *more efficient allocation of labour supply.* The redistribution achieved by the policy makes low-productivity households relatively wealthier, and the associated wealth effects reduce their labour supply.[29] The opposite occurs with high-productivity households. These changes result in a significant increase in average labour productivity—measured by the ratio of effective labour to hours worked—which can be seen in Figure 4(f). 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 favour of some amount of redistribution.

## 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 5(a). This translates into less overall risk and inequality. To show this in a compact way, it is useful to define a consumption–leisure composite, $c^{\gamma}(1-h)^{1-\gamma}$, which is the term that enters the households' period utility function. In Figure 5(b) and (c), we show that the optimal policy implies a reduction in risk (measured by the variance of the growth rate of the composite) that households face, and a reduction in the amount of inequality (measured by the Gini coefficient of the composite).

The reduction in inequality of the composite, however, masks a different effect of the policy on consumption and hours. Figure 5(d) and (e) shows that the policy implies a significant reduction in consumption inequality, but an increase in hours inequality. This increase in hours inequality is associated with the more efficient allocation of labour supply highlighted above.

## 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.[30]

### 5.3.1. **Average welfare gains.**

Consider a policy reform and denote by $\{c_t^j, h_t^j\}_{t=0}^{\infty}$ the equilibrium consumption and labour paths of a household with and without the reform, with

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

30. In Supplementary Appendix E.3, we consider an alternative decomposition that aims at setting apart the effects of policy on consumption and labour-supply decisions. We also present there decomposition results conditional of income and wealth quantiles.

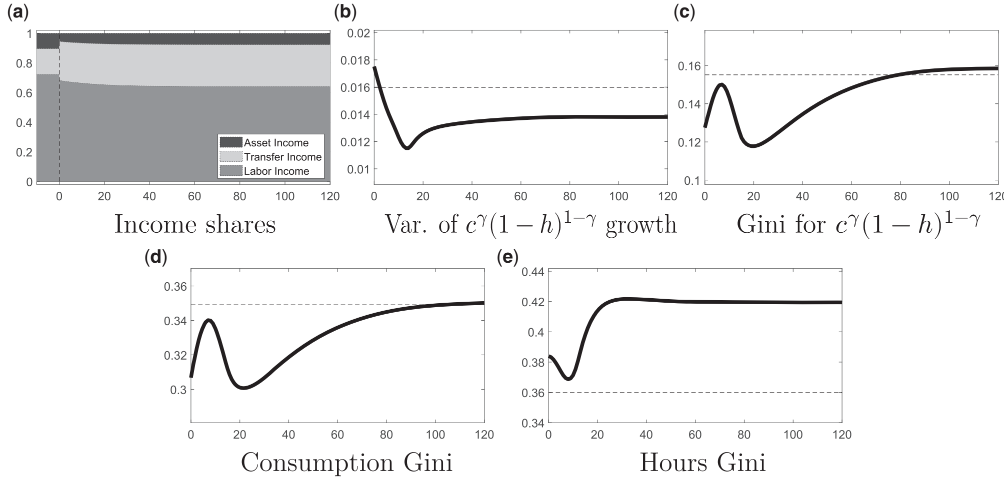


FIGURE 5

Optimal fiscal policy: distributional effects

*Notes:* (b)–(e): Thin dashed lines: initial stationary equilibrium; Thick solid curves: optimal transition.

$j=R$ or $j=NR$, respectively. The average welfare gain, $\Delta$, 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,

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

where $\lambda_0$ 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 $\lambda_t^j$ is the distribution over $(a_0,e^t)$ conditional on whether or not the reform is implemented with $e^t$ denoting the history of productivity realizations from period 0 to $t$. The level effect, $\Delta_L$, is then given by

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

**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 labour 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 $\{\bar{c}_t^j(a_0,e_0),\bar{h}_t^j(a_0,e_0)\}_{t=0}^{\infty}$ denote a certainty-equivalent sequence of consumption and labour 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 the associated 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, $\Delta_I$, is defined by

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

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

**3. Redistribution effect.** Utilitarian welfare also increases if the inequality across households with different initial states is reduced. A tax reform reduces 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, $\Delta_R$, can be defined as

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

Analogously to $p_{\text{risk}}^j$, $p_{\text{ineq}}^j$ 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 receive.

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

**Proposition 3.** *For balanced-growth-path preferences,*[31] *the components defined above satisfy the following relationship,*

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

Note that none of the elements of the decomposition are defined residually, hence this is indeed a decomposition and not a definition.

31. The proof in Supplementary Appendix E establishes the result for a more general set of utility functions.

TABLE 4
*Welfare decomposition for the benchmark and the fixed-instrument experiments*

| | | $\Delta$ | $\Delta_L$ | $\Delta_I$ | $\Delta_R$ |
|---|---|---|---|---|---|
| | **Benchmark** | 3.5 | 0.2 | 1.2 | 2.1 |
| **Instrument** | **Other instruments** | | | | |
| Fixed[a] capital income tax | Benchmark[b] | 0.8 | −0.6 | 1.3 | 0.1 |
| | Re-optimized[c] | 1.1 | −0.7 | 1.3 | 0.5 |
| Fixed labour income tax | Benchmark | 2.0 | 0.6 | −0.3 | 1.7 |
| | Re-optimized | 2.7 | 0.6 | 0.3 | 1.8 |
| Constant lump-sum[d] | Benchmark | 3.3 | −0.1 | 1.3 | 2.1 |
| | Re-optimized | 3.4 | 0.1 | 1.3 | 2.0 |
| Fixed lump-sum | Re-optimized | 2.1 | 1.0 | 0.0 | 1.0 |
| Fixed debt-to-output | Benchmark | 3.2 | −0.1 | 1.3 | 2.0 |
| | Re-optimized | 3.3 | 0.0 | 1.3 | 1.9 |

*Notes*: (a) "Fixed" means fixed at the initial stationary equilibrium value. (b) By "Benchmark" we mean keeping the other instruments at their benchmark optimal paths except for adjusting the level of lump-sum transfers to balance the intertemporal budget constraint of the government, so the economy is still in equilibrium. (c) In the "Re-optimized" experiments, we recompute the optimal path for the other instruments policy with the added restriction that one of the instruments is fixed. (d) In the "Constant lump-sum" experiments, we allow lump-sum transfers to move in period 0 but then restrict their path to be constant over time at that level.

**5.3.3. Choice of certainty equivalents.** There can be many certainty-equivalent paths that satisfy equation (5.3). These paths could differ over time and over levels of consumption and labour. In general, these choices can affect the components of the decomposition, but they are immaterial if household certainty equivalents follow parallel paths over time.

**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* $\{\tilde{C}_t^j\}_{t=0}^{\infty}$, *and* $\{\tilde{H}_t^j\}_{t=0}^{\infty}$.

Under this assumption, which we discuss in detail in Supplementary Appendix E, 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* $\Delta_L$, $\Delta_I$, *and* $\Delta_R$ *are independent of the paths* $\{\tilde{C}_t^j\}_{t=0}^{\infty}$, *and* $\{\tilde{H}_t^j\}_{t=0}^{\infty}$.

All welfare-decomposition results we present were calculated using certainty-equivalent paths that satisfy Assumption 1.

**5.3.4. Results.** The first row of Table 4 shows the welfare decomposition for our benchmark results. The optimal policy generates average welfare gains, $\Delta$, of 3.5%. Almost two-thirds of these gains, 2.1%, can be attributed to the redistribution effect, $\Delta_R$. The insurance effect, $\Delta_I$, implies an additional 1.2%, and the level effect, $\Delta_L$, captures the remaining 0.2% of gains. The experiments in the next two subsections are designed to shed light on how each fiscal instrument contributes to these gains.

### 5.4. *Fixed instruments*

To help clarify the role played by each instrument in the optimal policy, Table 4 also presents results for fixed-instrument experiments in which we hold each instrument fixed at their level in the initial stationary equilibrium. We present two versions of these experiments that are complementary. In the first version, for each fixed instrument, we simply set all other instruments to their benchmark optimal paths. We want the economy to still be in equilibrium though, so we adjust the level of lump-sum transfers to balance the intertemporal budget constraint of the government. In the second version, we re-optimize all other instruments while adding the fixed instrument restriction as a constraint for the planner.[32] For lump-sum transfers, we re-optimize under the constraint that they are constant over time while being able to move in Period 0, and under the constraint that lump-sum cannot move at all and is simply fixed in its initial steady-state level.

**5.4.1. Capital income taxes.** Changes to capital income taxes are the key source of the redistributive gains implied by the optimal policy. This is made clear by the fact that, regardless of whether or not we re-optimize the other instruments, fixing capital income taxes at their initial steady-state level leads to a substantial reduction in these gains. Perhaps more surprising, is the also substantial drop in the level effect. This is mostly due to the loss of average labour productivity improvements that result from redistribution. We return to this point in Section 6.

**5.4.2. Labour income taxes.** The second most welfare-relevant instrument is the labour income tax. Fixing it at its pre-reform level reduces average welfare by roughly 1.5% without re-optimization and 0.8% if the other instruments are re-optimized. Most of the welfare losses are associated with the insurance channel. The increase in the level component of welfare, highlights the relevant trade-off as more insurance comes at the cost of more distortions to labour supply decisions. Notice that the results so far are exactly in line with what we found in the two period example from Section 2: capital income taxes play a key role in the provision of redistribution, while changes in the labour income taxes are most important for the provision of insurance.

**5.4.3. Lump-sum transfers.** We conduct two types of experiments with lump-sum transfers. In the first, which we refer to by "Constant lump-sum" experiments, we allow lump-sum transfers to move in period 0 but then it must remain at that level in all future periods. This experiment shows that the optimal time variation of lump-sum transfers has small welfare implications relative to the optimal once-and-for-all increase.[33] When other instruments are re-optimized, the corresponding welfare losses are of about 0.1%. This indicates that the reasons behind the optimal time-varying lump-sum path are subtle. We return to this issue in the next subsection. In the "Fixed lump-sum" experiment, we set lump-sum transfers to be equal to their pre-reform levels in every period and re-optimize the other instruments. The average welfare gains are, in this case, reduced by 1.4%. When lump-sum transfers are not allowed to move, the planner provides redistribution by reducing labour income taxes. Since the labour income of

32. Supplementary Appendices O.11 and O.12 contain the figures for the re-optimized instruments and their associated aggregates.

33. Since time-variation of lump-sum transfers is not particularly important, one way to implement the overall increase in transfers in our model would be with the introduction of a constant universal basic income. To get a sense of magnitude, the increase of transfers is equivalent to 6% of GDP or 327 dollars a month (using 2019 GDP per capita in current prices). There is an increasing literature evaluating the benefits of UBI, see Guner *et al.* (2021), Luduvice (2019), and Daruich and Fernández (2020).

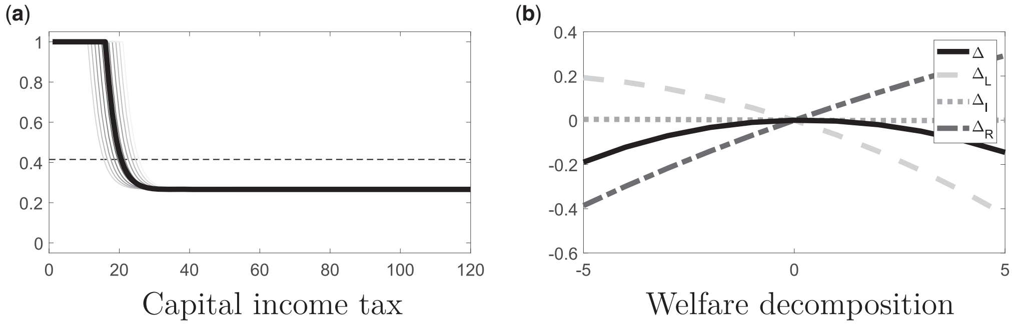


FIGURE 6

Varying the number of years capital income taxes are kept at the upper bound

*Notes:* (a) Thin dashed line: initial stationary equilibrium; Thick solid curve: optimal transition; Thin shaded solid curves: perturbations in the number of years capital income tax hits the upper bound from $-5$ to $+5$ relative to benchmark (b) the $x$-axis represents the change in the number of periods capital income taxes are kept at the upper bound relative to the optimum, $y$-axis shows change in the welfare gains in percent points.

lower productivity households is relatively low the amount of redistribution obtained is reduced by about half. The lower labour income taxes also imply that insurance gains disappear, while the level effect is improved. It follows that the overall increase in lump-sum transfers in the optimal benchmark policy plays a crucial role in the amount of redistribution and insurance implied by that policy.

**5.4.4. Debt-to-output.** Fixing the government debt-to-output ratio at the initial level reduces average welfare gains by 0.3% without re-optimization and by 0.2% when other instruments are re-optimized. In a similar way to what happens in the "Constant lump-sum" experiment, the majority of these relatively small losses come from the level effect. This is indicative of the fact that variations in government debt, as well as the timing of lump-sum transfers, allow the planner to mitigate the distortions associated with capital and labour income taxes. In the next subsection, we argue that this mitigation is achieved mostly by the effect of these instruments on the proportion of households that are borrowing constrained.

## 5.5. *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 economic mechanisms driving the optimal paths. 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.

**5.5.1. 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%. Figure 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, which can be appreciated by noticing that changing the number of years of capital income confiscation up or down by 5 years leads to average welfare changes of less than 0.2%.

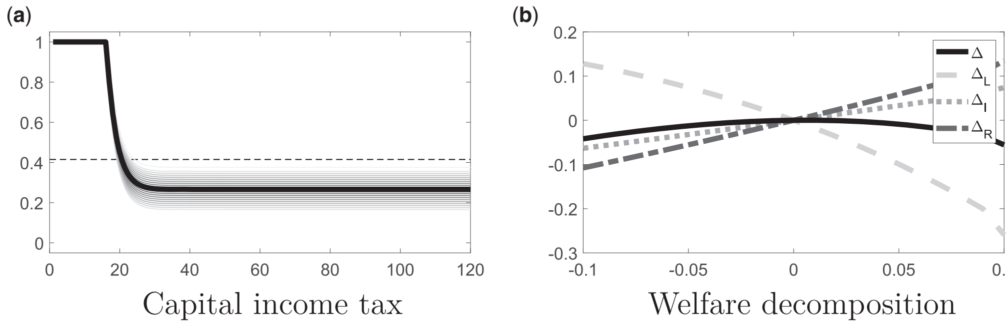


FIGURE 7

Varying long-run capital income taxes

*Notes:* (a) Thin dashed line: initial stationary equilibrium; Thick solid curve: optimal transition; Thin shaded solid curves: perturbations in the rate of capital income taxes starting from period 16 onward, from −10 to +10% relative to benchmark; (b) The *x*-axis represents the change in long-run capital income taxes relative to the optimum, *y*-axis shows change in the welfare gains in percent points.

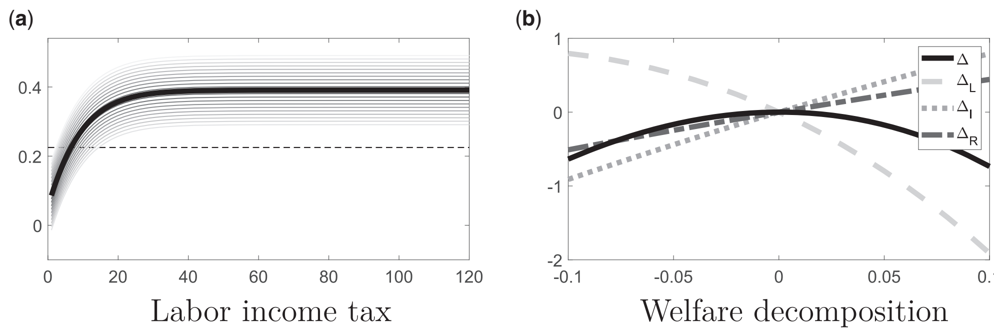


FIGURE 8

Varying labour income taxes

*Notes:* (a) Thin dashed line: initial stationary equilibrium; Thick solid curve: optimal transition; Thin shaded solid curves: perturbations in the rate of labour income taxes, from −10 to +10% relative to benchmark; (b) The *x*-axis represents the change in labour income taxes relative to the optimum, *y*-axis shows change in the welfare gains in percent points.

**5.5.2. 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.

**5.5.3. Labour income taxes.** Here, we change the average level of labour income taxes up and down by 10 percentage points, leading to the results in Figure 8. First notice that the effect of changes in labour 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

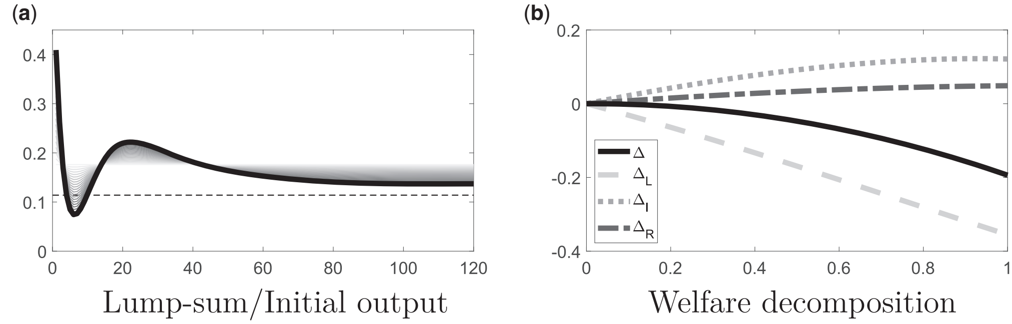


FIGURE 9

Varying lump-sum transfers

*Notes:* (a) Thin dashed line: initial stationary equilibrium; Thick solid curve: optimal transition; Thin shaded solid curves: perturbations towards constant lump-sum transfers; (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.

larger than the redistribution effect. Hence, though labour 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 labour income tax which is rebated via lump-sum transfers (exactly the experiment here) effectively reduces the labour income risk to which households are exposed.

**5.5.4. 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%.[34] 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.[35] 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 substantial 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.[36]

## 5.6. *Long-run optimality conditions*

Aiyagari (1995) analyses optimal long-run capital income taxes in an environment similar to ours. He argues that the Ramsey planner's decision to move aggregate resources across time

34. Notice that it does not follow from this that changes to the timing of lump-sum transfers *cannot* have important welfare implications. In Supplementary Appendix H.3, we show that backloading lump-sum transfers increases the share of borrowing-constrained households which can significantly reduce welfare.

35. 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 labour 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.

36. We illustrate these effects in Supplementary Appendix H, which provides more details about the perturbation towards constant transfers and an additional perturbation towards a monotonically decreasing path for lump-sum transfers.

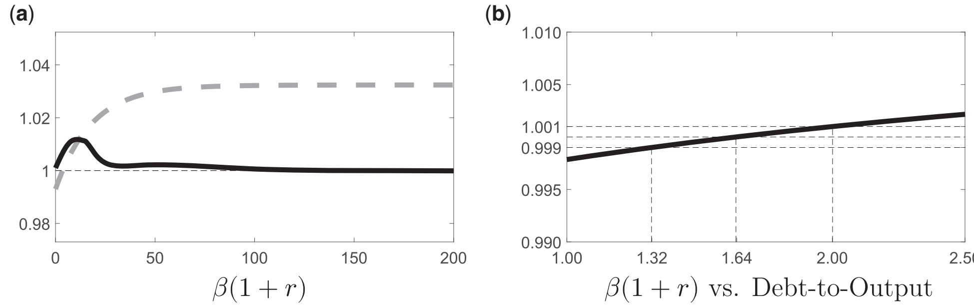


FIGURE 10

MGR and debt sensitivity.

*Notes:* (a) Thin dashed line: constant equal to one; Thick solid curve: $\beta(1+r)$ over time in the benchmark experiment; Thick dashed curve: optimal transition with constant policy (see Section 7.1). (b) The $x$-axis displays different levels of long-run debt-to-output; Horizontal thin dashed lines: 0.999, 1.000, 1.001; Thick solid curve: $\beta(1+r)$.

is risk-free and the associated Euler equation, in the long run, implies the MGR.[37] Lining this up with households' precautionary motivation for savings rationalizes positive long-run capital income taxes. Figure 10(a) shows that the MGR is satisfied in our benchmark results. We view this as corroborating evidence for the accuracy of our numerical long-run results. This accuracy is fundamentally important for pinning down the long-run optimal policies. As we demonstrate in Figure 10(b), and discuss extensively in Supplementary Appendix M, small ($\pm 0.1\%$) deviations from the MGR lead to large variations in the long-run debt-to-output ratio (from 1.32 to 2.00).

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 towards it and show that the MGR and three additional optimality conditions must hold. In Supplementary Appendix M, we extend their results to the balanced-growth-path preferences used in this article 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. Quantitative differences between our results and theirs must, therefore, be due to differences in the calibration and not the solution method. In Supplementary Appendix M, we also compare the two calibrations and discuss in detail the likely roots of these differences.[38] 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

37. The proof in Aiyagari (1995) that the MGR is a long-run optimality condition depends crucially on government spending being endogenous in his model, entering separately into the utility function of households. Acikgoz *et al.* (2018) show that the result generalizes to environments without endogenous government spending.

38. The most stark differences are that they find substantially higher optimal labour income taxes and debt-to-output ratios than we do. The higher levels of labour income taxes result, to a large extent, from stronger wealth effects on labour supply under their calibration. Supplementary Appendix M presents a detailed comparison between the two calibrations and how, in particular, our strategy leads to a significantly better fit to the distributions of earnings, wealth, and hours worked which also indirectly discipline wealth effects.

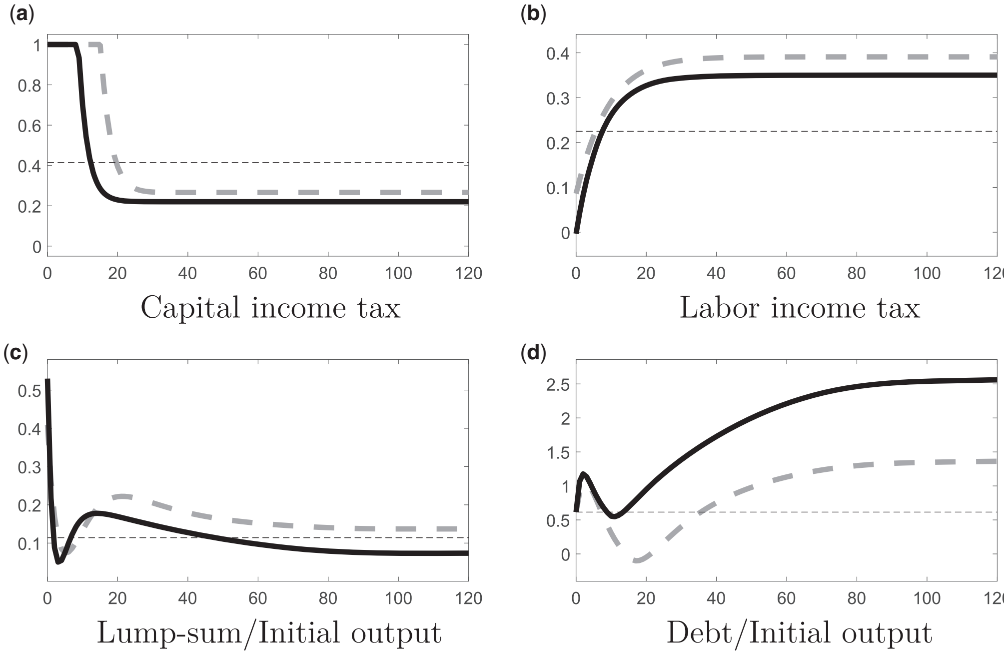


FIGURE 11

Optimal fiscal policy: maximizing efficiency

*Notes:* Thin dashed lines: initial stationary equilibrium; Thick solid curves: path that maximizes efficiency optimal transition; Thick dashed curves: path that maximizes the utilitarian welfare function (benchmark results).

equality-vs.-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 $\lambda_0$ 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 $W^{\sigma}$ is equivalent to maximizing the utilitarian welfare function. If $\hat{\sigma} \to \infty$, this becomes the Rawlsian welfare function. Finally, if $\hat{\sigma} = 0$, maximizing $W^0$ is equivalent to maximizing efficiency, where by *efficiency* we mean the combination of the level and insurance effects. We formalize claim in the following proposition.

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

In Supplementary 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 $\hat{\sigma} = 0$.[39] Figure 11 presents the results in comparison with the benchmark results. Relative

39. 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.

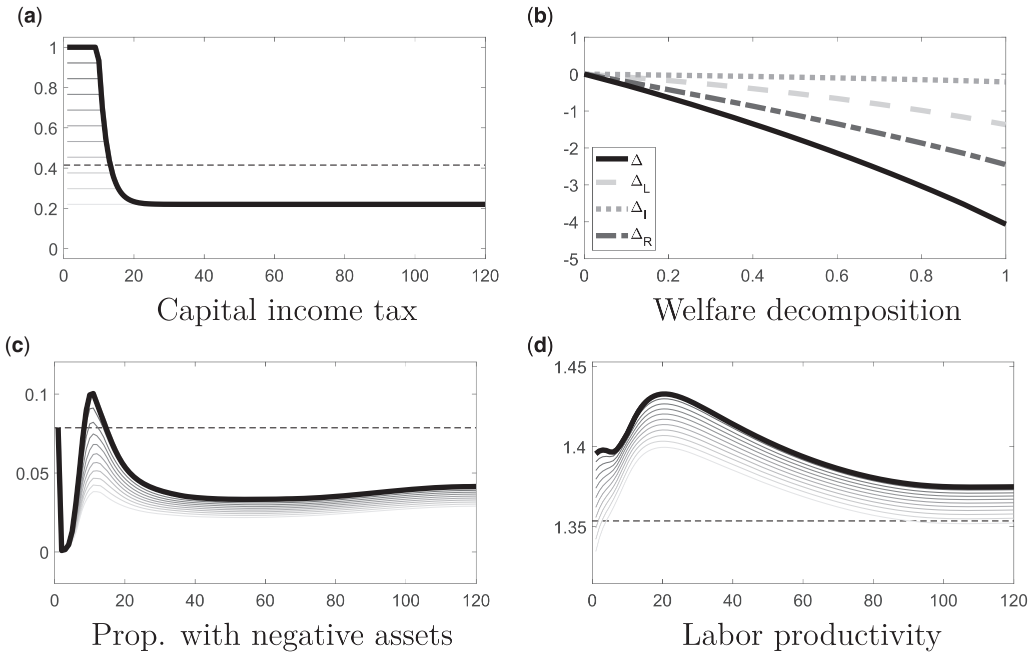


FIGURE 12

Reducing initial capital income taxes

*Notes:* (a,c,d) Thin dashed lines: initial stationary equilibrium; Thick solid curves: path that maximizes efficiency; Thin shaded solid curves: variations associated with the reduction in the initial capital income taxes; (b) the $x$-axis represents the homotopy parameter between the initial optimal path at $x=0$ and a constant capital income tax path at $x=1$, $y$-axis shows change in the welfare gains in percentage points.

to the initial stationary equilibrium, the policy implies average welfare gains of 1.8%: 0.8% from reduction in distortions and 1.0% from extra insurance. Even though the planner does not take this into consideration, the policy also implies a redistributive gain of about 1.1%.[40]

Relative to the benchmark experiment, capital and labour 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.

### 6.1. *Redistribution leads to efficiency gains.*

It is not at all obvious why it is optimal, with the purpose of maximizing efficiency, to confiscate capital income for the first 8 years. In a representative-agent setup without lump-sum taxes, the reason for front-loading capital 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 transfers in every period by the same amount to balance the budget.

First, notice from Figure 12(b) 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

40. Supplementary Appendix O.2 contains the figures for aggregates associated with this experiment.

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 labour productivity, which follows from the redistribution achieved by the high initial capital income taxes. As explained above, redistribution generates wealth effects on labour supply that lead to a more efficient allocation of hours in the economy, with higher productivity households working relatively more—see Figure 12(d). This effect is strong enough that it outweighs the distortions associated with the high initial capital taxes.[41]

### 6.2. *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 Supplementary Appendix I, 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.[42] 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. Finally, the optimal path of lump-sum transfers is monotonically decreasing in this case. This is indicative of the fact that the non-monotonicities found in the benchmark experiment are associated with capital income taxes staying at the upper bound for several periods before converging to a constant in the long run.

## 7. IMPORTANCE OF TIME-VARYING POLICIES

In this section, we illustrate the importance of allowing policy instruments to vary over time. As a first step to solve the Ramsey problem, we solved for the optimal once-and-for-all policy in which the planner must keep policy instruments constant after an initial change. We, then, proceeded by adding flexibility to our approximation in the time domain until we reach the benchmark approximation. We summarize some stages of this process in Table 5, and in Figures 13 and 14.

### 7.1. *Constant policy*

As can be seen in Figure 13,[43] the optimal once-and-for-all policy is essentially a weighted average of the time-varying instruments from our benchmark results. More weight is put on the short-run levels since those periods are more relevant for welfare. The long-run levels of the fiscal instruments differ substantially. Therefore, if one is interested in the long-run properties of the fiscal instruments, it is important to allow them to vary over time. In particular, as we noticed

41. This effect is not present in an earlier version of this article, Dyrda and Pedroni (2016), because there we assume a utility function without wealth effects on labour supply.

42. 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.

43. Figures with the corresponding aggregates are presented in Supplementary Appendix O.3.

TABLE 5
*Effects of time-varying policy*

| | $\tau^k$ | $\tau^h$ | $T/Y$ | $B/Y$ | $K/Y$ | $\Delta$ | $\Delta_L$ | $\Delta_I$ | $\Delta_R$ |
|---|---|---|---|---|---|---|---|---|---|
| Initial equilibrium | 41.5 | 22.5 | 11.4 | 61.5 | 2.49 | – | – | – | – |
| Constant policy | 67.5 | 27.9 | 19.7 | 53.9 | 2.02 | 1.6 | −0.7 | 0.8 | 1.6 |
| Front-loading | 54.7 | 29.4 | 18.9 | −1.0 | 2.36 | 2.8 | −0.3 | 0.8 | 2.3 |
| More flexibility (8 par.) | 34.4 | 40.2 | 21.2 | 29.2 | 2.49 | 3.4 | 0.1 | 1.3 | 2.1 |
| **Benchmark (17 par.)** | 26.7 | 39.1 | 15.2 | 154.3 | 2.48 | 3.5 | 0.2 | 1.2 | 2.1 |

*Notes:* All values, except for $K/Y$, are in percentage points. For $\tau^k$, $\tau^h$, $T/Y$, $B/Y$, and $K/Y$ in rows 2 to 5 we report values in the final stationary equilibrium. The average welfare, $\Delta$, and its components, $\Delta_L$, $\Delta_I$, and $\Delta_R$, are computed accounting for transition.

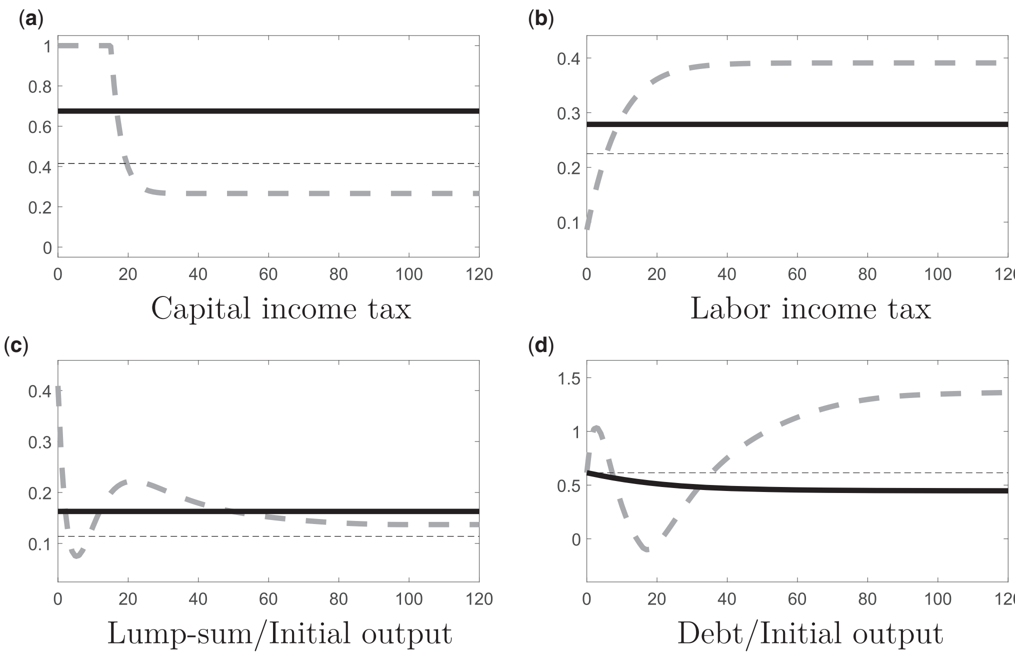


FIGURE 13
Optimal fiscal policy: constant policy

*Notes:* Thin dashed lines: initial stationary equilibrium; Thick solid curves: path that maximizes efficiency optimal transition; Thick dashed curves: benchmark results.

above in Section 5.6, whereas the MGR holds for the benchmark policy, it does not hold under the constant-policy restriction—see Figure 10(a). Moreover, constant policy leads to welfare gains that are less than half those of the optimal dynamic policy, as can be seen by comparing the second and last rows of Table 5. This difference in welfare is driven mostly by the level effect, which imply losses of −0.7 for constant policy and gains of 0.2 for the benchmark policy. This is indicative of the fact that time variation of fiscal instruments is important for the cross-mitigation of distortions. For instance, the initial paths of labour income taxes and lump-sum transfers help mitigate the distortions associated with high capital income taxes in the initial periods, something that is ruled out in the constant-policy experiment.

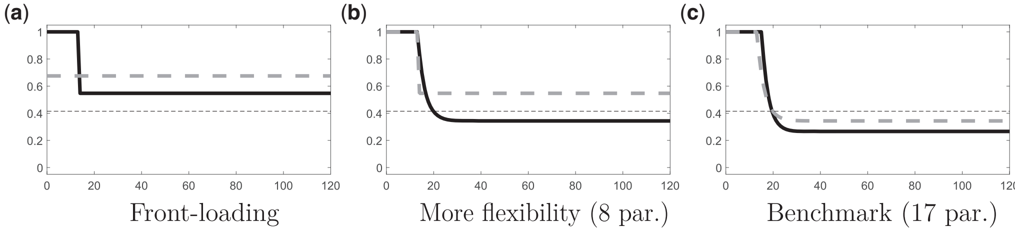


FIGURE 14

Adding flexibility to paths: capital income taxes

*Notes:* Thin dashed lines: initial stationary equilibrium; Thick dashed curve in (a): optimal constant taxes; Thick solid curve in (a) and thick dashed curve in (b): optimal transition allowing front-loading of capital income taxes; Thick solid curve in (b) and thick dashed curve in (c): optimal transition with eight parameters ($\alpha_0^k, \beta_0^k, \lambda^k, \alpha_0^h, \beta_0^h, \lambda^h, \beta_0^T, \lambda^T$); Thick solid curve in (c): benchmark optimal transition with 17 parameters—using $m_{\tau^k F} = m_{\tau^h F} = 0$, $m_{\tau^k 0} = m_{\tau^h 0} = m_{\tau^h F} = m_{TF} = 2$, and $m_{T0} = 4$ in equation (3.2).

### 7.2. *Front-loading capital income taxes*

In Figure 14, we focus on the path for capital income taxes, but at each stage all fiscal instruments are re-optimized. Figure 14(a) shows what happens when we allow capital income taxes to be front-loaded: this minimal amount of flexibility already increases welfare gains from 1.6% to 2.8, as reported in Table 5. Front-loading implies a substantial increase in the redistribution component of welfare, from 1.6 to 2.3%. It also improves the level effect by 0.4%, due to the more efficient allocation of labour implied by the additional redistribution.

### 7.3. *More flexibility (eight parameters)*

In Figure 14(b), we show what happens to capital income taxes when all fiscal instruments are allowed to follow the simplest form of equation (3.2), with polynomials of degree zero. This involves choosing eight parameters and the corresponding optimal policy improves welfare gains to 3.4%. Finally, Figure 14(c) shows what happens when we move from the 8-parameter solution to our benchmark 17-parameter solution, which brings welfare gains to 3.5%. The benchmark solution trades off a reduction in the insurance gains (from 1.26 to 1.19) for a more than offsetting increase in the level effect (from 0.05 to 0.23), while maintaining the redistributive gains—see the last three columns of Table 5. These results underscore that fine-tuning the time-variation of fiscal instruments can have important implications for what is achieved with the optimal policy.

In Supplementary Appendix G.3, we document all the additional intermediate steps of our implementation of this procedure with the corresponding figures and welfare gains. At each step in which we add more flexibility, welfare increases by less, but some of the fiscal instruments still change in meaningful. These changes compound to the differences in long-run instruments that can be observed between the fourth and last rows of Table 5. So, to determine optimal long-run policy accurately we make sure to keep adding flexibility until both welfare *and* policy are no longer affected.

## 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, labour-income and wealth inequality, in turn. Introducing uninsurable idiosyncratic productivity risk and borrowing constraints brings us back to the SIM model. At each step, we analyse 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, $\Gamma$, set to the identity matrix and borrowing constraints replaced by no-Ponzi conditions. In order to keep the amount of labour-income inequality comparable with the benchmark calibration, we rescale the productivity levels so as to keep the variance of the present value of labour income the same. Since the wealth distribution is indeterminate in the steady state of this economy, as argued by Chatterjee (1994), we can set the initial distribution to be the same as in our benchmark economy. We recalibrate the discount factor, $\beta$, 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 $\Theta$ such that the optimal tax system is given by $\tau_t^k = 1$ for $0 \leq t < t^*$; while for $t \geq t^*$ $\tau_t^k$ 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 $0 \leq t \leq t^*$, $\tau_t^h$ evolves according to*

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

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

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

In Supplementary Appendix F, we apply the method introduced by Werning (2007) to prove this proposition,[44] and analogous ones for versions of this economy without labour–income and/or wealth inequality.[45] In particular, we also show that the magnitudes of $t^*$ and $\Theta$ are related to the levels of wealth and labour–income inequality, respectively. Figure 15 illustrates the numerical results obtained using this proposition.

44. 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.

45. In the economy without wealth inequality, lump-sum transfers and capital income taxes in Period 0 are non-distortive and have no effect on redistribution, so their optimal levels are indeterminate.

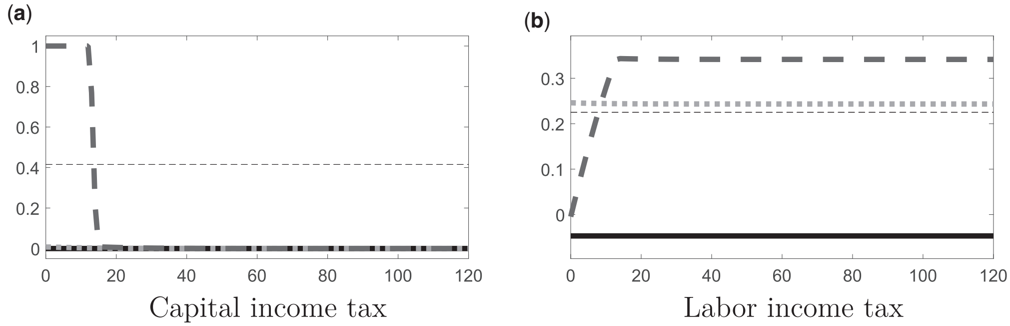


FIGURE 15

Optimal taxes: complete market economies

*Notes:* Thin dashed lines: initial taxes; Thick solid curves: optimal taxes for representative economy; Thick dotted curves: optimal taxes with only labour-income inequality; Thick dashed curve: optimal taxes with labour-income and wealth inequality.

### 8.1. *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 labour 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 labour income taxes set equal to the negative of the consumption tax ensures labour supply decisions are not distorted either.

### 8.2. *Labour-income inequality.*

When labour 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 labour income taxes be virtually constant over time and capital income taxes virtually equal to zero in every period.

### 8.3. *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.[46] While capital income taxes are at the upper bound, labour income taxes are increasing. This leads to a decreasing (or less increasing) path for labour supply, which mitigates distortions to the households' intertemporal decisions: it leads to a smoother path for period utility as leisure increases while consumption decreases.

### 8.4. *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

46. 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 Supplementary Appendix F.8, we include a more detailed discussion of this issue.

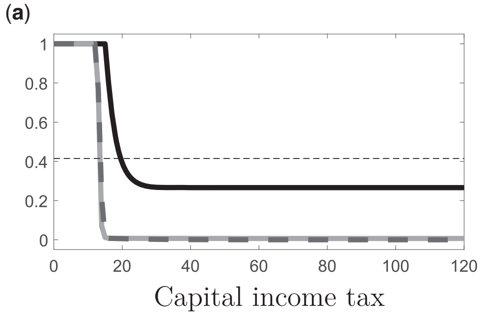


FIGURE 16

Optimal taxes: complete market economies

*Notes*: Thin dashed lines: initial taxes; Thick solid curves: optimal taxes from Benchmark SIM model; Thick solid shaded curves: optimal taxes calculated using the same parameterized paths used in the Benchmark experiment; Thick dashed curves: optimal taxes calculated using Proposition 6.

for this economy, the parameterized paths are able to approximated the actual solution relatively well (average welfare gains are similar as well: 2.253% using the proposition vs. 2.246% 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.

## 9. SENSITIVITY ANALYSIS AND ROBUSTNESS

In Supplementary 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 and they show up mostly on the number of periods capital income taxes remain in the upper bound: which is reduced to 10 years for an IES of 0.8, and increased to 71 years for an IES of 0.5. Finally, we show that increases in the Frisch elasticity unsurprisingly reduce labour income taxes though by relatively small amounts.

In Supplementary Appendices M and N, we present results for four alternative calibrations: (1) an economy that disciplines the labour 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: (1) the qualitative features of the Ramsey policy in the SIM model that we highlight in the article—high short-run capital income taxes combined with increasing labour income taxes—are robust to substantial changes to the calibration; (2) the quantitative results are sensitive to the calibration, which justifies the extensive effort we put into all details of it.

## 10. CONCLUDING COMMENTS

In this article, we quantitatively characterize the solution to the Ramsey problem in the SIM 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%. 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 labour via wealth effects on labour supply. Increasing labour income taxes over time and a non-monotonic path for lump-sum transfers mitigate the intertemporal distortions associated with high capital income taxes. Government debt has relatively small welfare consequences, in part because, for the majority of the optimal transition, only a minority of households are borrowing constrained, but also because the associated general equilibrium price effects have counteracting effects on redistribution and insurance.

Finally, this article 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 article 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.

*Acknowledgments*. We are grateful to Dirk Krueger, the Editor, as well as to four anonymous referees for careful readings and helpful recommendations. We also 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. 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.

**Supplementary Data**

Supplementary data are available at *Review of Economic Studies* online. And the replication packages are available at https://dx.doi.org/10.5281/zenodo.6462485.

**Data Availability Statement**

The data underlying this article are available at https://doi.org/10.5281/zenodo.6462485.

REFERENCES

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

ACIKGOZ, O., HAGEDORN, M., HOLTER, H. and WANG, Y. (2018), "The Optimum Quantity of Capital and Debt" (Technical Report).

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 MCGRATTAN, E. R. (1998), "The Optimum Quantity of Debt", *Journal of Monetary Economics*, **42**, 447–469.

ATKESON, A., CHARI, V. and KEHOE, P. J. (1999), "Taxing Capital Income: A Bad Idea", *Federal Reserve Bank of Minneapolis Quarterly Review*, **23**, 3–18.

BAKIS, O., KAYMAK, B. and POSCHKE, M. (2015), "Transitional Dynamics and the Optimal Progressivity of Income Redistribution", *Review of Economic Dynamics*, **18**, 679–693.

BANSAL, R., KIKU, D. and YARON, A. (2012), "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices", *Critical Finance Review*, **1**, 183–221.

BANSAL, R. and YARON, A. (2004), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles", *The Journal of Finance*, **59**, 1481–1509.

BARRO, R. J. (2009), "Rare Disasters, Asset Prices, and Welfare Costs", *American Economic Review*, **99**, 243–264.

BARSKY, R. B., MANKIW, N. G. and ZELDES, S. P. (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 Hildenbrand, W. and Mas-Collel, A. (eds) *Contributions to Mathematical Economics in Honor of Gerard Debreu* (Amsterdam: North-Holland).

BHANDARI, A., EVANS, D., GOLOSOV, M. and SARGENT, T. J. (2017), "Public Debt in Economies with Heterogeneous Agents", *Journal of Monetary Economics*, **91**, 39–51.

BOAR, C. and MIDRIGAN, V. (2020), "Efficient Redistribution" (Working Paper 27622, National Bureau of Economic Research).

CASTAÑEDA, A., DÍAZ-GIMÉNEZ, J. and RÍOS-RULL, J.-V. (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., NICOLINI, J. P. and TELES, P. (2018), "Optimal Capital Taxation Revisited" (Working Papers 752, Federal Reserve Bank of Minneapolis).

CHATTERJEE, S. (1994), "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model", *Journal of Public Economics*, **54**, 97–119.

CHEN, Y., YANG, C. C. and CHIEN, Y. (2020), "Implementing the Modified Golden Rule? Optimal Ramsey Capital Taxation with Incomplete Markets Revisited" (Working Papers 2017-003, Federal Reserve Bank of St. Louis).

CHETTY, R. (2006), "A New Method of Estimating Risk Aversion", *American Economic Review*, **96**, 1821–1834.

CHETTY, R., GUREN, A., MANOLI, D. and WEBER, A. (2011), "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins", *American Economic Review*, **101**, 471–475.

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

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

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

DARUICH, D. and FERNÁNDEZ, R. (2020), "Universal Basic Income: A Dynamic Assessment" (NBER Working Papers 27351, National Bureau of Economic Research, Inc.).

DÁVILA, J., HONG, J. H., KRUSELL, P. and RÍOS-RULL, J.-V. (2012), "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks", *Econometrica*, **80**, 2431–2467.

DÍAZ-GIMÉNEZ, J., GLOVER, A. and RÍOS-RULL, J.-V. (2011), "Facts on the Distributions of Earnings, Income, and Wealth in the United States: 2007 Update", *Federal Reserve Bank of Minneapolis Quarterly Review*, **34**, 2–31.

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

DYRDA, S. and PEDRONI, M. Z. (2016), "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Shocks" (2016 Meeting Papers 1245, Society for Economic Dynamics).

EROSA, A., FUSTER, L. and KAMBOUROV, G. (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., KAJII, A. and NAKAJIMA, T. (2015), "Optimal Taxation and Debt with Uninsurable Risks to Human Capital Accumulation", *American Economic Review*, **105**, 3443–3470.

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

GREULICH, K., LACZÓ, S. and MARCET, A. (2019), "Pareto-Improving Optimal Capital and Labor Taxes" (Technical Report).

GRUBER, J. (2013), "A Tax-Based Estimate of the Elasticity of Intertemporal Substitution", *The Quarterly Journal of Finance*, **3**, 1350001.

GUNER, N., KAYGUSUZ, R. and VENTURA, G. (2021), "Rethinking the Welfare State" (Technical Report, Centre for Economic Policy Research).

GUVENEN, F. (2011), "Macroeconomics with Heterogeneity: A Practical Guide", *Economic Quarterly*, **97**, 255–326.

HANSEN, L. P., HEATON, J., LEE, J. and ROUSSANOV, N. (2007), "Intertemporal Substitution and Risk Aversion", *Handbook of Econometrics*, **6**, 3967–4056.

HEATHCOTE, J., PERRI, F. and VIOLANTE, G. L. (2010), "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006", *Review of Economic Dynamics*, **13**, 15–51.

HEATHCOTE, J., STORESLETTEN, K. and VIOLANTE, G. L. (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**, 1364–1383.

ITSKHOKI, O. and MOLL, B. (2019), "Optimal Development Policies with Financial Frictions", *Econometrica*, **87**, 139–173.

JONES, L. E., MANUELLI, R. E. and ROSSI, P. E. (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 TIMMER, G. T. (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 KRUEGER, D. (2021), "High Marginal Tax Rates on the Top 1 Percent? Lessons from a Life-Cycle Model with Idiosyncratic Income Risk", *American Economic Journal: Macroeconomics*, **14**, 319–366.

KRUEGER, D. and LUDWIG, A. (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).

LUDUVICE, A. V. D. (2019), "The Macroeconomic Effects of Universal Basic Income Programs" (Working Paper No. 202121, Federal Reserve Bank of Cleveland).

MARCET, A., OBIOLS-HOMS, F. and WEIL, P. (2007), "Incomplete Markets, Labor Supply and Capital Accumulation", *Journal of Monetary Economics*, **54**, 2621–2635.

NAKAJIMA, M. and RÍOS-RULL, J.-V. (2019), "Credit, Bankruptcy, and Aggregate Fluctuations" (Working Papers 19-48, Federal Reserve Bank of Philadelphia).

NUÑO, G. and THOMAS, C. (2016), "Optimal Monetary Policy with Heterogeneous Agents (Updated September 2019)" (Working Papers 1624, Banco de España; Working Papers Homepage).

PONCE, M., VAN ZON, R., NORTHRUP, S., GRUNER, D., CHEN, J., ERTINAZ, F., FEDOSEEV, A., GROER, L., MAO, F., MUNDIM, B. C., NOLTA, M., PINTO, J., SALDARRIAGA, M., SLAVNIC, V., SPENCE, E., YU, C.-H. and PELTIER, W. R. (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), PEARC '19* (New York, NY, USA: Association for Computing Machinery).

POWELL, M. (2009), "The BOBYQA Algorithm for Bound Constrained Optimization without Derivatives" (Technical Report, Department of Applied Mathematics and Theoretical Physics, Cambridge University).

PRUITT, S. and TURNER, N. (2020), "Earnings Risk in the Household: Evidence from Millions of US Tax Returns", *American Economic Review: Insights*, **2**, 237–254.

RAGOT, X. and GRAND, F. L. (2020), "Optimal Fiscal Policy with Heterogeneous Agents and Aggregate Shocks" (2017 Meeting Papers 969, Society for Economic Dynamics).

RÖHRS, S. and WINTER, C. (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 Cooley, T. F. (ed.) *Frontiers of Business Cycle Research* (Princeton, NJ: Princeton University Press) 294–330.

SAEZ, E. and STANTCHEVA, S. (2018), "A Simpler Theory of Optimal Capital Taxation", *Journal of Public Economics*, **162**, 120–142.

STRAUB, L. and WERNING, I. (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 UHLIG, H. (2012), "How Do Laffer Curves Differ across Countries?" in Alesina, A. and Giavazzi F. (ed.) *Fiscal Policy after the Financial Crisis* (Chicago: University of Chicago Press) 211–249.

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