Optimal Public Debt with Redistribution [*]

Matias Bayas-Erazo[†]

This version: July 19, 2026
First version: November 21, 2023

Abstract

Governments choose both the progressivity of the tax system and the amount of public debt in circulation. How should the two be set together? I study fully dynamic optimal policy in standard heterogeneous-agent incomplete-markets economies, solving for both the long-run optimum and the transition towards it. In the long run, I find that planners with stronger preferences for redistribution choose more progressive tax systems but *lower* levels of public debt. This is driven by an interest rate channel: redistributive taxation reduces the demand for safe assets and makes government borrowing more expensive. Along the transition, the same channel operates in reverse. The government initially flattens the tax schedule, and the resulting fall in interest rates helps finance lower taxes and higher consumption in the short run.

Keywords: public debt, redistribution, heterogeneous-agent models, optimal fiscal policy

[*]This paper is a revised version of Chapter 1 of my PhD dissertation at Northwestern University. I am grateful to my advisors Guido Lorenzoni, George-Marios Angeletos, Alessandro Pavan, and Matthew Rognlie for their insights, guidance, and support throughout this project. I also thank Martin Beraja, Anmol Bhandari, Corina Boar, Rachel Childers, Axelle Ferriere, Joao Guerreiro, Fergal Hanks, Christian Hellwig, David Hémous, Diego Huerta, Felix Kubler, Xavier Ragot, Florian Scheuer, Ludwig Straub, Nicolas Werquin, Yucheng Yang, and Kwok Yan Chiu for comments that helped improve the paper. All remaining errors are my own.

[†]University of Zürich. Email: matias.bayas-erazo@econ.uzh.ch.

1 Introduction

There are ongoing debates on how public debt and progressive income taxes should be used. Several economists and policy makers have argued that, when interest rates are low, governments can and should borrow more (see **?**). At the same time, progressive taxation is seen by many as a tool to address the recent increase in income and wealth inequality (**?**; **?**). While there is extensive research on each of these instruments *in isolation*, we know less about how debt and progressivity should be set together. The two are connected, however: the interest rate at which the government borrows depends, in equilibrium, on how much insurance the tax system provides. How much insurance the tax system can provide depends, in turn, on how much the government borrows.

When markets are incomplete, both public debt and progressive income taxes can help households insure against risk. Public debt provides liquidity and contributes to the supply of safe assets that agents use to self-insure (**?**; **?**). Meanwhile, a progressive income tax transfers resources from the lucky to the unlucky and acts as a form of social insurance (**?**; **?**). If countries choose a progressive tax system due to a high demand for insurance, then it makes sense for them to also favor higher levels of public debt. On the other hand, if progressive taxation lowers the aggregate demand for safe assets, it may increase the cost of borrowing, perhaps weakening the incentive to issue public debt. What, then, is the optimal *mix* of debt and progressivity, and how does it depend on social preferences for redistribution?

To address these questions, I compute the optimal mix of debt and redistributive taxation in standard heterogeneous-agent incomplete-markets models (**?**; **?**; **?**). Households face uninsurable idiosyncratic income risk, supply labor, and self-insure by holding safe assets subject to borrowing constraints. A perfectly competitive firm hires labor and produces final output. The government controls the supply of safe assets and a progressive tax on labor income. In this environment, I study fully dynamic optimal policy and solve for both the long-run optimum and the transition towards it.

In the long run, I find that planners with stronger preferences for redistribution implement more progressive tax systems, as one would expect. But, perhaps surprisingly, they also favor *lower* levels of public debt. In other words, differences in preferences for redistribution generate

a negative correlation between progressivity and indebtedness.

To explain what is behind this result, I derive a simple condition that characterizes the optimal long-run level of public debt in these economies. Public debt provides liquidity, but servicing it is costly. The condition organizes this tradeoff into three key "sufficient statistics": i) the marginal social value of public debt, ii) the sensitivity of interest rates to changes in the level of public debt, and iii) the premium on public debt. This premium, the difference between the discount rate and the interest rate, arises due to incomplete markets. Together, these three objects determine the optimal level of public debt in the long run.

Increasing the progressivity of the tax system puts upward pressure on interest rates, thereby lowering the premium on public debt. This *interest rate channel of progressivity* arises because a progressive tax system reduces the need to self-insure against idiosyncratic income risk. Agents are willing to pay less for holding safe assets, and government borrowing becomes more expensive. At the same time, a progressive tax system reduces the benefit households derive from public liquidity. When the government is already offering insurance through the tax system, the value of self-insurance goes down, reducing the marginal social value of public debt. Both effects push in the same direction: more redistribution raises the cost of issuing debt and weakens one of its benefits. As a result, an inequality-averse planner, who naturally favors a redistributive tax system, finds it optimal to issue lower levels of public debt.

When public debt is not part of the analysis, studies typically find that, under a utilitarian criterion, the tax system should be more progressive than what we observe in the US (**?**). The same is true in the model when debt is fixed at the observed level. Once the level of debt is chosen together with the tax system, this is no longer the case. In fact, the planners closest to the utilitarian benchmark choose regressive tax schedules, well below the progressivity estimated for the US. The interest rate channel helps account for the difference: it pushes the planner to favor less progressivity once the level of debt is part of the problem.

So far, what varies is the planner's taste for redistribution, with the economic environment held fixed. A natural question is what happens when the environment varies instead. If economies differ in the amount of idiosyncratic risk their households face, the optimal mix moves in the opposite direction: economies with more risk choose more progressivity *and* more debt. This is because risk raises both the value of redistributive taxation and the demand for safe assets.

The relationship between debt and progressivity under optimal policy thus depends on why economies differ. This distinction matters when reading the data through the lens of the model. Among advanced economies, the association is negative, pointing to differences in preferences rather than in risk. Building on this, I later back out each country's implied preference for redistribution.

Computing optimal policy in this environment involves two technical challenges. The first is well known: the planner must keep track of a time-varying wealth distribution. To overcome this, I formulate the problem in sequence space, following the approach introduced by **?**. The second is that the long-run optimum—what the literature calls the Ramsey steady state—may not exist under standard welfare criteria. **?** study the Ramsey problem in a stylized economy that abstracts from heterogeneity and show that long-run satiation can be optimal. **?** find a similar pattern in the full model: under utilitarian welfare and separable preferences, the planner issues debt until the interest rate equals the discount rate. At that point, however, asset holdings diverge, which is inconsistent with a well-defined steady state.

Given this, I depart from utilitarian welfare criteria and consider planners that use generalized stochastic welfare weights (**?**). These weights are used to aggregate individual utilities and determine how much the planner values additional consumption for different agents. They are allowed to depend on endogenous outcomes and thus decouple society's concerns for fairness from agents' lifetime welfare. To the extent that infinitely-lived agents represent altruistically linked generations, these weights allow the planner to care about the inequality of future generations directly (**?**; **?**). I refer to such planners as generational planners. Although nonstandard, this class of welfare functions nests the utilitarian criterion and allows me to speak about inequality-aversion while keeping the problem tractable. I verify numerically that an interior Ramsey steady state exists for the inequality-averse specifications I consider: a planner that values redistribution from the asset-rich to the asset-poor no longer finds satiation optimal. This helps overcome some of the difficulties that previous work has faced in this class of models.

Leveraging this, I also study the optimal *transition* towards the Ramsey steady state. Computing this is more involved: the planner must choose policy instruments at each date along a potentially long horizon. I address this challenge by anchoring policy paths to the Ramsey

steady state and parameterizing deviations using a low-dimensional time basis. I then exploit the sequence-space structure to compute policy gradients analytically, combining insights from **?** and **?**. This makes the problem tractable on accessible hardware and delivers, to the best of my knowledge, the first computed optimal transition for the joint mix of debt and progressivity in this class of models.

Along the optimal transition, the interest rate channel continues to play a central role. The planner starts by flattening the tax schedule: progressivity falls and average taxes decline. Because the tax system now provides less insurance, the private demand for safe assets rises, pushing interest rates down. The resulting increase in the premium on public debt creates fiscal space that helps finance the reform. This allows the planner to sustain lower taxes and higher consumption in the short run while gradually moving the debt-to-GDP ratio towards the long-run target.

To explore the robustness of these results, I consider several extensions. I first repeat the analysis ignoring transitional dynamics, computing the mix of debt and progressivity that maximizes steady-state welfare à la **?**. Because this problem is computationally more tractable, I can use more standard social welfare functions. In doing so, I verify that the properties of the optimal long-run mix are robust to the details behind the social welfare function, although the steady-state planner chooses lower levels of debt and higher levels of progressivity.

Then, I consider more flexible labor income tax schedules that allow for lump-sum transfers, which are absent from the baseline. This changes the shape of the optimal tax schedule (**?**) but planners that care about redistribution still favor lower levels of public debt. The same is true in a version of the model with multiple safe assets and taxes on savings, where the negative correlation between progressivity and indebtedness becomes *stronger*.

A simple open economy variant, in which some demand for safe assets comes from abroad (**?**), provides a more direct test of the mechanism. When the interest rate is less responsive to domestic conditions, the negative correlation between optimal progressivity and public debt becomes weaker.

Finally, I turn the question around and carry out an inverse optimum exercise (**?**; **?**). I use each country's observed mix of debt and progressivity to recover its model-implied preference for redistribution. The resulting ranking is stable across specifications and it lines up with in-

dependent evidence on redistributive preferences: Japan and the US at one end, Germany and the Nordic economies at the other.

Related literature

This paper contributes to the literature on optimal fiscal policy with heterogeneous agents that begins with **?**. Assuming the existence of the Ramsey steady state, that paper shows that the long-run optimum features positive capital income taxes. Due to the difficulties involved in tracking the wealth distribution, most studies deviate from the original Ramsey problem and follow the seminal work of **?**, who compute the level of debt that maximizes steady-state welfare. Like most of the literature, they restrict attention to linear taxes and thus ignore the interaction with progressive taxation that is the focus of this paper. One notable exception is **?**, who studies the optimal steady-state mix of debt and transfers. He also finds that debt may provide insurance more effectively whenever equity considerations are not part of the welfare objective. However, he does not relate the properties of the optimal mix to the interest rate channel, and more importantly, continues to ignore transitions.

? solve the full Ramsey problem in a stylized incomplete-markets economy that bypasses the computational challenges of the problem considered in this paper. They clarify how the approach taken by **?** and **?** ends up overestimating the costs of issuing public debt and thus underestimates its long-run quantity.

Relatedly, several papers in the quantitative Ramsey tradition point out that accounting for the transition path can lead to very different optimal tax schedules in environments with heterogeneous agents. **?** focus on once-and-for-all changes in the tax system and find that accounting for transitions leads to a more progressive optimal tax system. **?** look at the interaction between progressive taxation and education subsidies and also find that the optimal progressivity of the tax system depends on whether or not transitional dynamics are taken into account. **?** also study once-and-for-all reforms, evaluating welfare consequences along the transition. They find small welfare gains from enriching the set of instruments available to the planner. All of these papers abstract from public debt and are thus unable to relate the optimal level of debt to redistribution.

More recently, the literature has returned to the original Ramsey problem in **?**, developing

different approaches to address the computational challenges. **?** use a Lagrangian approach inspired by **?**. **?** directly search for the optimal sequence of policies after parameterizing them in the time domain. **?** introduce a sequence-space approach for computing the Ramsey steady state and find that the standard heterogeneous agent model with utilitarian welfare criteria has no interior steady state.[1] **?** focus on optimal policy over the business cycle and opt for an inverse optimal taxation approach, estimating a social welfare function that makes the current tax system consistent with the planner's optimality conditions. Relative to these papers, I extend the sequence-space approach of **?** to solve for both the Ramsey steady state and the optimal transition. To compute the transition, I parameterize policy paths in the time domain, as in **?**, but use the sequence-space representation to efficiently evaluate equilibrium responses. By anchoring to the RSS and using gradient-based optimization, my approach runs on accessible hardware.

Given the focus on redistribution, this paper also speaks to a literature that is closer to the Mirrleesian approach to optimal taxation. Authors working with static models tend to emphasize the equity-efficiency tradeoff, but **?** points out that redistributive taxation can also be viewed as a form of social insurance. He shows that a government can effectively insure individuals against income risk, a theme that is explored further by the literature working with dynamic Mirrleesian models (see **?** and **?**, among others). Unlike this paper, the Mirrleesian approach tends to work in partial equilibrium and disregards the role of public debt. An important exception is **?**, who studies nonlinear fiscal policy in a model with complete markets. He finds that the relationship between taxes and debt is indeterminate (i.e. Ricardian equivalence holds).[2] **?** look at nonlinear tax policy in the standard incomplete markets model but assume away the role of public debt by forcing the government to balance the budget every period. Similarly, **?** focus on the optimal design of transfers and progressivity.

[1] To get around the non-existence of the RSS with separable preferences, **?** rely on GHH preferences, whereas **?** use a KPR utility function.

[2] Relatedly, **?** show that Ricardian equivalence holds in the standard incomplete markets model when the government can also control the tightness of borrowing constraints. **?** consider fiscal policy and debt management jointly but restrict attention to proportional taxes.

2 Model

Consider a standard incomplete markets economy with a continuum of households who face uninsurable idiosyncratic income risk (**?**; **?**). The only asset is a one-period risk-free government bond that pays an interest rate r and can be freely traded up to some borrowing limit $\phi \geq 0$. Individual productivity θ evolves according to some Markov process and determines the wage per unit of labor supplied by the agents. There is no aggregate uncertainty.

Given a sequence of interest rates $\{r_t\}$ and nonlinear labor income tax schedules $\{T_t(\cdot)\}$, individuals face an income fluctuation problem with endogenous labor supply. The value function of an agent entering the period with assets a and productivity θ in period t is

$$V_t(a,\theta) = \max_{\ell,c,a'} u(c) - v(\ell) + \beta \mathbb{E}_{\theta'|\theta}\left[V_{t+1}(a',\theta')\right] \quad \text{s.t} \quad \begin{cases} c + a' = (1+r_t)a + \theta\ell - T_t(\theta\ell) \\ a' \geq -\phi, \end{cases}$$

where c, ℓ, and a' are consumption, labor supply, and next period's asset holdings and $\beta \in (0,1)$ is the agent's discount factor. Pre-tax labor income is given by $y = \theta\ell$. For future reference, let $\boldsymbol{c}_t(x)$, $\boldsymbol{a}_t(x)$, and $\boldsymbol{y}_t(x)$ denote the policy functions for consumption, asset holdings, and labor income for an agent in state $x = (a,\theta)$. Also, denote by $D_t(\theta, A)$ the measure of households with productivity θ that have assets in set A at the beginning of period t.

For the baseline, I restrict attention to tax schedules that exhibit a constant rate of progressivity (CRP), as in **?** and **?**:

$$T_t(y) = y - \tau_t \, y^{1-p_t}, \tag{1}$$

for some $\{p_t, \tau_t\}$ with $p_t < 1$ and $\tau_t > 0$, for all t. The parameter p_t indexes the progressivity of the tax schedule in period t, whereas τ_t governs the average level of after-tax wages. Taxes are linear if $p_t = 0$, progressive if $p_t > 0$, and regressive if $p_t < 0$. With this functional form, after-tax income $\boldsymbol{z}_t(x) = \tau_t \, \boldsymbol{y}_t(x)^{1-p_t}$.

Notice that this rules out the possibility of lump-sum transfers, an important feature of tax and transfer systems in practice. I start with this benchmark because it captures progressivity with a single parameter that has a direct empirical counterpart. I relax it in Section 5, where I consider alternative labor income tax schedules that allow for transfers.

To close the model, I assume that the government supplies the risk-free bond subject to a

standard budget constraint

$$G + r_{t-1}B_{t-1} = B_t - B_{t-1} + \int T_t(\boldsymbol{y}_t(x))dD_t(x), \quad \text{for all } t. \tag{2}$$

In words, the government's exogenous spending needs $G \geq 0$ and the interest payments on the debt must be financed by aggregate tax revenues and net debt issuance in period t. On the production side, I assume that the effective labor ($\theta\ell$) of different productivity types is perfectly substitutable in production. This means that each unit of effective labor produces one unit of goods and that the real wage is equal to one. Goods market clearing requires

$$\int \boldsymbol{y}_t(x)dD_t(x) = \int \boldsymbol{c}_t(x)dD_t(x) + G, \quad \text{for all } t. \tag{3}$$

Finally, the asset market clearing condition is

$$\int \boldsymbol{a}_t(x)dD_t(x) = B_t, \quad \text{for all } t. \tag{4}$$

Note that the only supply of safe assets outside the household sector comes from the government. In particular, there is no capital. I start with this assumption in order to isolate the role of public debt but I relax it in Section 5, where I consider a more general production technology and allow firms to issue claims to physical capital.

Definition 1 An equilibrium is a sequence of interest rates $\{r_t\}$ and tax schedules $\{T_t\}$, a sequence of policy functions $\{\boldsymbol{c}_t(x), \boldsymbol{a}_t(x), \boldsymbol{y}_t(x)\}$ and distributions $\{D_t\}$ such that, given an initial distribution D_0:

(i) $\{\boldsymbol{c}_t(x), \boldsymbol{a}_t(x), \boldsymbol{y}_t(x)\}$ are optimal given $\{r_t, T_t\}$,

(ii) D_t is consistent with the policy functions and the Markov process for productivity,

(iii) the government's budget constraint (2) is satisfied,

(iv) the asset market clearing condition (4) and the goods market condition (3) hold. □

2.1 Calibration

I calibrate the model to the US economy, following **?** whenever possible. A period is one quarter and there is no borrowing (i.e. $\phi = 0$). The discount factor β is chosen such that the ratio of aggregate liquid assets to GDP in the model is consistent with US data, given a real interest rate of 2%.[3] Following standard practice in the literature, I assume that θ follows an AR(1) process in logs with persistence ρ and an innovation variance σ_ϵ^2. These parameters are chosen to match **?**'s estimates for the persistence of the US wage process and the standard deviation of earnings growth in **?**. I discretize the AR(1) process for productivity using the Rouwenhorst method on eight idiosyncratic states. The elasticity of intertemporal substitution is equal to one and the Frisch elasticity of labor supply is 1/2, consistent with **?**. For taxes, the value for progressivity p is taken from **?**. Government spending is 8.8% of annual GDP, the average ratio of government expenditures to output in the US over the period 1970 to 2013.[4] Given these choices, the average after-tax wage τ is pinned down by the government's budget constraint. Table 1 summarizes the parameters that come out from this procedure.

Table 1: Parameter values

Parameter	*Description*	*Value*	*Parameter*	*Description*	*Value*
β	discounting	0.986	G/Y	spending-to-GDP	0.088
ρ	persistence of AR (1)	0.966	B/Y	debt-to-GDP	1.4
σ_ϵ^2	shock variance	0.033	p	tax progressivity	0.181
EIS	curvature in u	1	τ	after-tax wage	0.912
Frisch	curvature in v	1/2	ϕ	borrowing limit	0

[3] **?** calculate liquid assets from aggregate household balance sheets reported in the flow of funds and take the average ratio over the period 1970 to 2013. They arrive at a value of $A/Y = 1.4$. Since bonds are the only asset, the model's debt-to-GDP ratio is higher than what is observed in the data.

[4] See Table 1.1.5. Gross Domestic Product in the National Income and Product Accounts.

3 Interest rate channel of progressivity

This section uses the model to study how progressivity affects the equilibrium interest rate. The analysis isolates the interest rate channel of progressivity, which drives the properties of the optimal mix. I show that progressive tax reforms raise equilibrium interest rates and that, as a result, they can have unintended distributional consequences that benefit the asset-rich.

Consider a small permanent change in the progressivity of the tax schedule dp around the baseline economy while holding the level of debt fixed. The goal is to understand how this reform affects the equilibrium interest rate, the level of taxes, and individual payoffs. For simplicity, I focus on stationary outcomes. Lemma 1 characterizes the response of these objects to this change in the tax system:

Lemma 1 *Given a small permanent change in progressivity dp,*

1. *The general equilibrium responses $d\boldsymbol{r}$ and $d\boldsymbol{\tau}$ are given by the solution to the following system:*

$$\begin{bmatrix} \frac{\partial \mathscr{A}^{ss}}{\partial r} & \frac{\partial \mathscr{A}^{ss}}{\partial \tau} \\ \frac{\partial \mathscr{T}^{ss}}{\partial r} - B & \frac{\partial \mathscr{T}^{ss}}{\partial \tau} \end{bmatrix} \begin{bmatrix} d\boldsymbol{r} \\ d\boldsymbol{\tau} \end{bmatrix} = \begin{bmatrix} -\frac{\partial \mathscr{A}^{ss}}{\partial p}\, dp \\ -\frac{\partial \mathscr{T}^{ss}}{\partial p}\, dp \end{bmatrix}.$$

2. *The response of individual outcomes dV solves:*

$$dV(x) = u'(\boldsymbol{c}(x)) \left(\boldsymbol{y}(x)^{1-p} d\boldsymbol{\tau} + a\, d\boldsymbol{r} - \boldsymbol{z}(x) \log \boldsymbol{y}(x)\, dp \right) + \beta \mathbb{E}_{\theta'|\theta} \left[dV(\boldsymbol{a}(x), \theta') \right]. \tag{5}$$

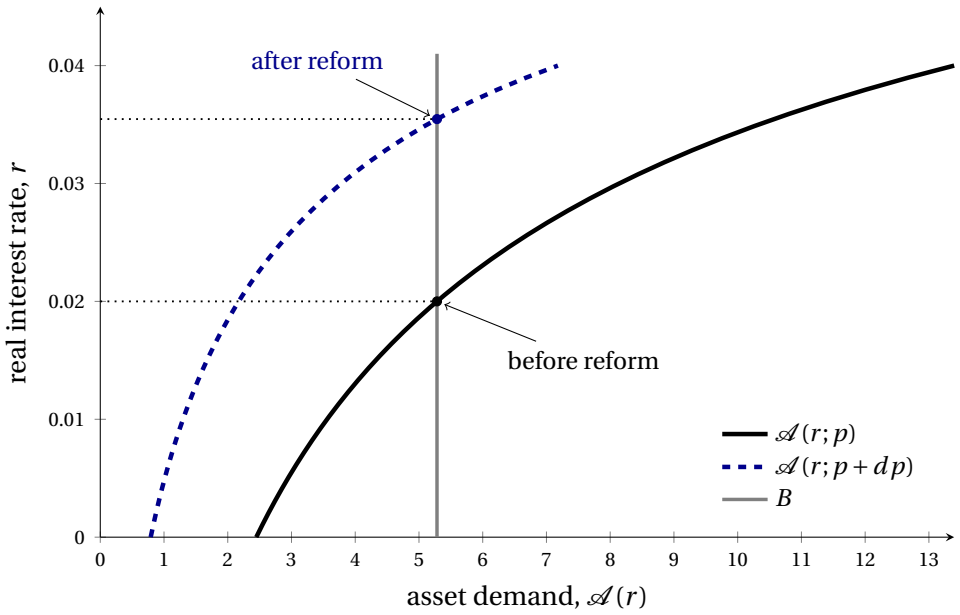
PROOF The first part follows from the implicit function theorem. The second part follows from the envelope theorem, taking into account general equilibrium effects. ∎

The interest rate response to the reform depends on the partial equilibrium responses of aggregate asset demand $\mathscr{A}^{ss}(r,p,\tau) = \int a_i\, di$ and aggregate tax revenues $\mathscr{T}^{ss}(r,p,\tau) = \int T(y_i)\, di$ to changes in interest rates and taxes around the initial steady state. Because these objects are functions of the distribution of agents over idiosyncratic states, it is not possible to characterize them analytically. However, a simple intuition explains the sign of the interest rate response.

For a given r, an increase in the progressivity of the tax system lowers the volatility of after-tax income and reduces aggregate asset demand. There is more insurance via the tax system

and thus less need to hold assets to self-insure against idiosyncratic income risk.[5] When the supply of debt is held fixed, the interest rate must rise in order to clear the asset market, so that $d\boldsymbol{r} > 0$. Figure 1 illustrates this mechanism by plotting the equilibrium in the asset market before and after the reform. The solid black curve is asset demand before the reform, the dashed blue curve is asset demand after the reform, and the grey line is the fixed supply of debt. The demand curve shifts to the left because the need for self-insurance is lower. With the supply of debt held fixed, the equilibrium interest rate rises. The difference between the two dotted lines is what I call the interest rate channel of progressivity.

Figure 1: Equilibrium in the asset market before and after progressive tax reform



One of the contributions of this paper is to trace out the implications of this mechanism for the optimal mix of debt and progressivity. A similar force appears in recent work by **?**, who show that inequality increases fiscal space in a two-agent model. Relatedly, **?** find that transfers reduce the demand for safe assets, which lowers the upper bound on deficits in a model with perpetual youth. **?** study how progressive tax reforms affect equilibrium interest rates in heterogeneous-agent incomplete-markets models and find results consistent with the mechanism I describe here. Similarly, **?** find that progressive tax systems reduce the maximum sus-

[5]There is an additional effect that works in the same direction: the increase in progressivity raises the level of taxes (i.e. $d\boldsymbol{\tau} < 0$). This lowers disposable income and thus shifts the demand for safe assets further to the left.

tainable deficit of the government. Taken together, their findings point to a potential conflict between progressive taxation and debt issuance. However, they do not study the optimal mix of these two instruments, which is the focus of this paper.

The quantitative relevance of the interest rate channel depends on the sensitivity of aggregate safe asset demand to changes in the real interest rate. Intuitively, if aggregate safe asset demand is sufficiently elastic, then a small change in interest rates suffices to clear the market. While direct empirical estimates of how tax progressivity affects interest rates are not available, there is substantial evidence on how public debt affects interest rates.[6] **?** survey this literature and report semi-elasticity estimates between 1.1% and 2.5%. My calibrated model yields an annualized semi-elasticity of 2.4%. This suggests the strength of the interest rate channel is in line with the empirical estimates.

The indirect general equilibrium effects of progressive tax reforms, operating through the interest rate response, can outweigh the direct partial equilibrium effects. As a result, progressive tax reforms may have unintended effects. The second part of Lemma 1 makes this concrete: a change in progressivity affects individual lifetime utilities through a *direct* and an *indirect* effect. The direct effect is due to the change in the slope of the tax schedule, which lowers taxes paid at the bottom of the income distribution and raises them at the top. This tends to favor the income-poor. The indirect effect operates through the general equilibrium responses of interest rates and the level of taxes, favoring those who are asset-rich at the time of the reform.

In the baseline calibration, the interest rate response is strong enough that the asset-rich *benefit* from progressive tax reforms, regardless of where they stand in the productivity distribution. On the other hand, those with low wealth are hurt by the reform, unless their productivity is low enough for the direct effect to dominate. Appendix A decomposes these effects across the state space.

Next, I show that the interest rate channel also shapes the optimal mix of debt and progressivity and helps explain why inequality-averse planners favor relatively low levels of public debt.

[6] **?** find that countries that are more unequal save less. This is not necessarily inconsistent with the mechanism I describe here: their cross-sectional evidence does not condition on financial development, and their own theory attributes the pattern to deeper financial markets in unequal countries, which let households borrow more.

4 Optimal policy

This section turns to optimal policy. The planner chooses the entire path of debt and progressivity, taking transitional dynamics into account. Building on **?**, I first derive a simple characterization of the Ramsey steady state and verify that an interior one exists for the planners I consider. I then characterize the optimal long-run mix of debt and progressivity and compute the transition towards it.

4.1 Ramsey problem

The Ramsey planner chooses *sequences* of interest rates, debt levels, and CRP tax codes to maximize the present discounted value of aggregate utility subject to implementability conditions. As mentioned earlier, this problem takes into account the transition path and allows all instruments to vary over time. I begin by characterizing the limiting steady state of the Ramsey plan. This reduces the dimensionality of the problem and provides a target for the transition analysis that follows.

Formally, the dynamic Ramsey problem is

$$\max_{\{r_t,B_t,p_t,\tau_t\}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) \text{ s.t } \begin{cases} \mathscr{A}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) = B_t, \\ G + (1+r_{t-1})B_{t-1} = B_t + \mathscr{T}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right). \end{cases} \tag{6}$$

Here, $\mathscr{U}_t(\cdot)$ is a sequence-space function that maps sequences of interest rates and tax schedules into "aggregate utility" at time t. Similarly, $\mathscr{A}_t(\cdot)$ and $\mathscr{T}_t(\cdot)$ map sequences of interest rates and taxes into aggregate asset holdings and aggregate tax revenues in period t. These functions aggregate optimal individual decisions using the distribution of agents in the economy. This ensures that the fiscal policy chosen by the planner can be implemented in equilibrium, provided that the government's budget constraint holds and the asset market clears. I assume that the Ramsey planner aggregates individual utilities as follows:

$$\mathscr{U}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) = \int_i \omega_t(a_{it})\{u(c_{it}) - v(\ell_{it})\}\, di, \tag{7}$$

with the weights $\omega_t(a) \propto \exp(-\alpha_a a)$.[7] The parameter $\alpha_a \in \mathbb{R}$ governs the planner's aversion to

[7] The weights depend on t because they are normalized by $\bar{\omega}_t = \int_i \exp(-\alpha_a a_{it})\, di$, the average weight in each period. This normalization ensures that the SWF does not penalize increases in asset holdings *per se*.

wealth inequality. This nests the utilitarian criterion when $\alpha_a = 0$. When $\alpha_a > 0$, the planner discounts the instantaneous utility of the asset-rich relative to a utilitarian benchmark.

This social welfare function (SWF) departs from standard welfare criteria for two reasons. First, it takes as input individual instantaneous utilities, as opposed to individual lifetime utilities. Second, it depends on weights that are endogenous. Formally, the planner uses generalized dynamic stochastic weights to conduct welfare assessments, in the sense of **?**. These weights allow one to formalize new welfare criteria that society may find appealing and decouple society's concerns for fairness from individual utilities, as in **?**. For example, making the weights decline with assets captures the view that society considers it fair to redistribute from the asset-rich to the asset-poor.

My preferred interpretation is that the social planner distinguishes between the welfare of each generation. To the extent that a model with infinitely lived agents represents altruistically linked generations of finitely lived agents, this SWF allows the planner to care *directly* about inequality within future generations (**?**; **?**).[8] In this interpretation, $\alpha_a > 0$ means that agents with relatively high assets receive less weight, while $\omega(a) = 1$ gives equal weight to all agents within a generation. I refer to planners with this welfare criterion as generational planners.

Allowing for departures from a utilitarian welfare criterion helps resolve the complications that arise when searching for the steady state of the Ramsey plan. The literature continues to debate the existence of an interior Ramsey steady state in the standard heterogeneous agent model. **?** claim that an interior Ramsey steady state does not exist when the elasticity of intertemporal substitution is less than one. In an environment with deterministic income fluctuations, **?** prove that the steady state exists for separable CRRA utility functions provided that the planner is not utilitarian. In a more standard calibration of the model, the results in **?** suggest that a version of the Friedman rule may be optimal when the Ramsey planner uses a utilitarian welfare criterion. In other words, a utilitarian planner finds it optimal to satiate the demand for public debt, issuing debt to the point where the interest rate is equal to the discount rate. However, this is not consistent with an interior steady state in this class of models: aggregate

[8] I do not want to take a stand on the philosophical debate of whether the welfare of future generations should enter the planner's objective. I am simply arguing that *if* society cares about the inequality of future generations, then the SWF would resemble (7).

asset demand asymptotes to infinity as $\beta(1+r) \to 1$.

In what follows, I derive the condition that characterizes the optimal long-run value of public debt and verify that, for the inequality-averse specifications I consider, it has a unique interior solution. A planner that puts less weight on the asset-rich assigns a lower marginal social value to public debt and no longer finds satiation optimal. Appendix B.5 contains the details and shows that the result continues to hold when debt and progressivity are chosen jointly. When debt is held fixed, the restricted problem over tax progressivity has an interior solution even for a utilitarian planner.

4.2 Sufficient statistics for the optimal long-run choice of debt

To solve the Ramsey problem, I rely on a set of first order conditions that I obtain by perturbing the optimal plan. Here, I derive a simple "sufficient-statistic" representation for the optimal long-run choice of debt. Appendix B derives the full set of first-order conditions for the Ramsey problem and elaborates on the computational approach to compute the steady state.

Fix an arbitrary sequence of CRP tax codes $\{p_t, \tau_t\}$. For any $u = 0, 1, \ldots$, the following must be true:

$$\sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial B_u} + \sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\beta^{t-u}\lambda_t\frac{\partial \mathscr{T}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial B_u} + \lambda_u - \beta\lambda_{u+1}(1+\boldsymbol{r}_u) - \sum_{t=0}^{\infty}\beta^{t-u}\lambda_t\frac{\partial \boldsymbol{r}_t}{\partial B_u}B_{t-1} = 0. \tag{8}$$

Here, λ_t denotes the Lagrange multiplier on the government's budget constraint in period t and $\boldsymbol{r}_t(\cdot)$ is a sequence-space function that maps sequences of CRP tax codes $\{\tau_t, p_t\}$ and public debt $\{B_t\}$ into the interest rate that clears the asset market at time t. If (8) does not hold, then a small perturbation dB_u would increase social welfare, contradicting the optimality of the Ramsey plan.

To understand the intuition behind each term, let us first ignore the summations. The first term says that increasing the amount of debt at time u has a direct effect on welfare via interest rates. The strength of this effect is mediated by the sensitivity of interest rates to changes in the level of public debt. The second term arises due to income effects on labor supply. The third term captures the effects on the government's budget: λ_u is the shadow value of public funds in period u, and issuing debt relaxes the budget at time u. However, this debt has to be repaid next period at the given interest rate, which explains the term $-\beta\lambda_{u+1}(1+\boldsymbol{r}_u)$. Finally, because

interest rates are endogenous in this model, the change in debt issuance is also affects interest rate payments on the existing debt.

Now, the double summations are there due to the forward/backward nature of the system. Individuals are forward looking and changes in the distribution propagate slowly. Therefore, welfare, tax revenues, and interest rates depend on the entire sequence of fiscal policy chosen by the government.

Despite this, Proposition 1 shows that this condition simplifies around the steady state:

Proposition 1 *Suppose that* $\lim_{u\to\infty}\lambda_u = \lambda^{RSS} > 0$. *Then, the optimal long-run level of debt* B^{RSS}*, if it exists, solves*

$$\left[\frac{\mathscr{E}_r^{\mathscr{U}}}{\lambda^{RSS}} + \mathscr{E}_r^{\mathscr{T}}\right]\mathscr{E}_B^{\boldsymbol{r}} + \{1-\beta(1+\boldsymbol{r})\} - \mathscr{E}_B^{\boldsymbol{r}}\, B^{RSS} = 0, \tag{9}$$

where $\mathscr{E}_X^F \equiv \lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial F_t}{\partial X_u}$.

PROOF See Appendix B.4. ∎

From this result, it is clear that the optimal long-run level of public debt can be written in terms of three key "sufficient statistics". First, the marginal social value of public debt, $\frac{\mathscr{E}_r^{\mathscr{U}}}{\lambda^{RSS}} + \mathscr{E}_r^{\mathscr{T}}$. The first component is the discounted long-run response of aggregate utility to changes in the interest rate, in dollar terms, while the second captures the associated response of tax revenues through income effects on labor supply.

Second, the so-called premium on public debt, which is the difference between the discount rate and the interest rate: $1-\beta(1+\boldsymbol{r})$. This arises due to incomplete markets and captures the rents that the government extracts from the private sector due to its ability to issue risk-free debt.

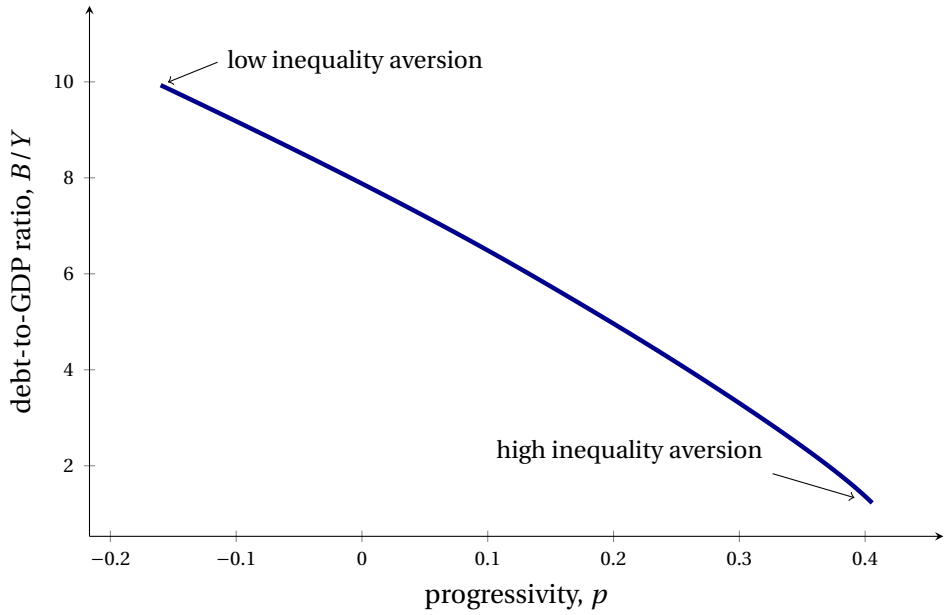
Finally, the sensitivity of interest rates to changes in public debt: $\mathscr{E}_B^{\boldsymbol{r}}$. This object captures the extent to which the government can affect the *path* of interest rates by issuing more or less debt. In models with complete markets, $\mathscr{E}_B^{\boldsymbol{r}} = 0$ and the condition reduces to $1 = \beta(1+r)$. Since this is already guaranteed by the consumer's Euler equation in steady state, the optimal level of public debt is indeterminate. In models with incomplete markets, $\mathscr{E}_B^{\boldsymbol{r}} \neq 0$ and the optimal level of public debt is determined by the relative strength of the three terms.

Equation (9) is quite general, and its logic extends beyond the specific structure of the model considered in this paper. It applies to a broad class of incomplete-markets economies, including those in the spirit of **?**, where debt acts as collateral, and overlapping-generations models, where debt acts as a vehicle for life-cycle savings (**?**). What changes across these environments is the mapping from fiscal policy choices to the sufficient statistics, but the relevant "reduced-form" objects are the ones identified in Proposition 1.

4.3 Optimal long-run mix of debt and progressivity

I now turn to the optimal long-run mix of debt and progressivity. Figure 2 plots the optimal mix for different values of α_a, the planner's aversion to wealth inequality. As α_a rises, the optimal tax system becomes more progressive and the optimal level of public debt falls.

Figure 2: Optimal long-run mix of debt and progressivity



To see why, consider the planners in the bottom-right corner of the figure, which are highly averse to wealth inequality. They naturally prefer a progressive tax system because it redistributes towards the agents they care about the most. Given the positive correlation between asset holdings and labor income, a progressive labor income tax helps the asset-poor. Now, because of the interest rate channel discussed in Section 3, this means they face relatively higher

interest rates in the long run: redistributive taxation reduces the need to self-insure against income risk. In order to induce agents to hold a given supply of safe assets, the government must eventually offer a higher interest rate. Hence, it is optimal to issue lower levels of public debt in the long run. To put it in terms of the sufficient statistics isolated in Proposition 1, the premium on public debt is lower when the planner favors a redistributive tax system.

In addition, one component of the value of issuing public debt falls when the need for self-insurance is lower. When the government provides more insurance through the tax system, $\mathcal{E}_r^{\mathcal{U}}$, which captures the response of aggregate utility to changes in interest rates, drops. This is a key determinant of the marginal social value of public debt, as shown in Proposition 1.[9] The combination of these two effects explains the negative correlation between progressivity and indebtedness.

This result contradicts the common presumption that governments with a stronger taste for redistribution should also borrow more. Redistribution and borrowing are usually seen as parts of the same package: a larger government that taxes more, transfers more, and borrows more. Here, a stronger taste for redistribution is precisely what makes borrowing more expensive.

Solving for debt and progressivity jointly also matters quantitatively. Taking the observed level of debt as given, a utilitarian planner chooses a progressivity of 0.375, above the 0.181 estimated for the US and consistent with the literature that abstracts from public debt (**?**). Once the level of debt is part of the problem, this is no longer the case. Although there is no interior Ramsey steady state under the utilitarian criterion, the planners on the frontier closest to this benchmark choose *regressive* tax schedules (i.e. $p < 0$).

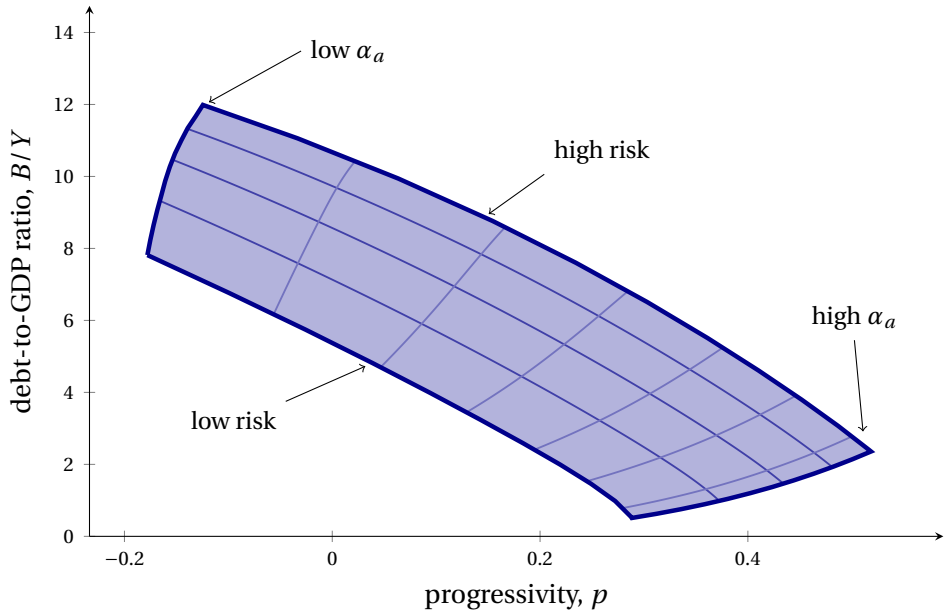
4.4 Preferences, risk, and the cross-section

So far, what varies along the frontier is the planner's taste for redistribution, with the economic environment held fixed. A natural question is what happens when the primitives vary instead. Figure 3 traces out the optimal long-run mix when the amount of idiosyncratic income risk varies as well. Holding risk fixed, planners with a stronger taste for redistribution choose more progressivity and less debt, as before. Holding the welfare criterion fixed, economies with more

[9] The overall effect on the marginal social value of public debt might be ambiguous, though, because income effects cause $\mathcal{E}_r^{\mathcal{T}}$ to increase with higher progressivity.

risk choose more progressivity *and* more debt: risk raises both the value of redistributive taxation and the demand for safe assets, so the two instruments now move together. The relationship between debt and progressivity under optimal policy thus depends on why economies differ. I use this below to read the cross-country evidence, and in Section 6 to recover each country's preference for redistribution from its observed mix.

Figure 3: Optimal long-run mix as inequality aversion and idiosyncratic income risk vary



Note: The underlying grid varies α_a from 0.0035 to 0.04 and the unconditional standard deviation of log productivity from 70 to 150 percent of its baseline value.

In the entire cross section, countries with more progressive tax systems tend to have, if anything, higher levels of public debt (Appendix F). This is the sign that variation in risk would generate, although the association weakens once differences in development, a force outside the model, are taken into account. More "similar" economies, where these differences are smaller, offer a more informative comparison. Among the economies that were already advanced in 1996, the association turns negative, and more so within the G7, though the samples are small.[10] Through the lens of the model, it is as if these economies differed mainly in their preferences

[10] Progressivity here is measured using taxes and transfers, the concept closest to the model. Appendix F reports both this measure and one based on income taxes alone, which moves the same way as the sample narrows, and shows that controlling for differences in government spending does not change the patterns.

for redistribution.

4.5 Optimal transition to the Ramsey steady state

With the optimal long-run policy in hand, I now turn to the optimal transition: given an initial distribution, what is the optimal time path of fiscal instruments that takes the economy to the Ramsey steady state?

The Ramsey problem (6) optimizes over sequences of policy variables. The computational challenge is that the planner chooses policy instruments at each date along a potentially long horizon. As described in the introduction, I address this by anchoring policy paths to the Ramsey steady state and parameterizing deviations using a low-dimensional time basis. The spirit of the approach is to combine the sequence-space formulation of **?** with the direct parameterization method of **?**. The implementability constraints are handled via an augmented Lagrangian:

$$\mathscr{L}_{\rho_{AL}}(x,\lambda) = -W(x) + \lambda^\top \boldsymbol{g}(x) + \frac{\rho_{AL}}{2}\|\boldsymbol{g}(x)\|^2, \tag{10}$$

where $x = \{r_t, p_t, \tau_t\}_{t=0}^T$ stacks the policy variables over the transition horizon, $W(x) = \sum_{t=0}^T \beta^t \mathscr{U}_t(x)$ is the discounted sum of aggregate utility, $\boldsymbol{g}(x)$ collects the goods market clearing residuals at each date[11], λ is the vector of Lagrange multipliers, and $\rho_{AL} > 0$ is a penalty parameter.

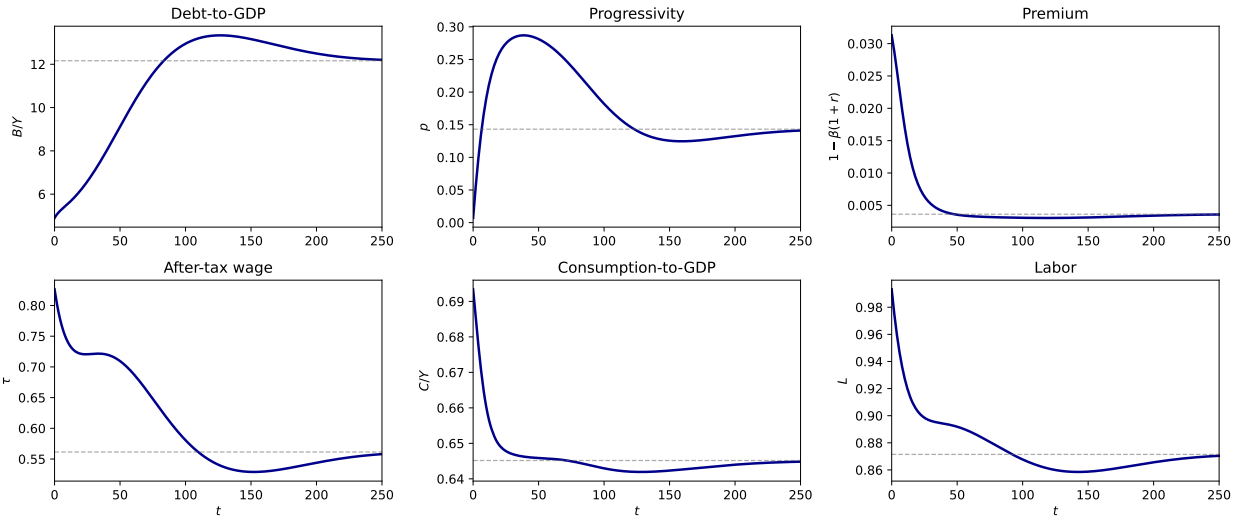
I solve this problem using policy gradient techniques, exploiting the sequence-space structure to compute the gradients analytically. The gradients of (10) involve sequence-space Jacobians around *arbitrary paths*.[12] To reduce the computational burden, I parallelize the construction of these Jacobians and use the recursive structure of the household problem to speed up the computation of each column. The details are in Appendix E.

Figure 4 shows the optimal transition paths for an inequality-averse planner ($\alpha_a = 0.025$) starting from the baseline calibration in Section 2. The planner starts by flattening the tax schedule: progressivity falls on impact and the average tax burden declines. This reduction in distortions triggers an immediate boom in output and consumption.

[11] By Walras' Law, the original problem (6) can be reformulated with goods market clearing as the only constraint, after dropping the government budget constraint and the associated choice variable B_t. See Appendix B.3 for details.

[12] This prevents the use of the efficient algorithm in **?**.

Figure 4: Optimal transition paths



Note: The transition horizon is $T_{\text{free}} = 425$ quarters using a Chebyshev time basis of degree 5 per instrument and an anchoring function $e_t = (1 - t/T_{\text{free}})^{\kappa}$ with $\kappa = 10$ that anchors the path to the Ramsey steady state. The transition appends a constant tail of length $T_{\text{tail}} = 375$ at terminal policies to approximate the infinite-horizon continuation. Dashed lines indicate Ramsey steady-state values.

The interest rate channel continues to operate along the transition, now in reverse. Because the tax system provides less insurance, the private demand for safe assets rises, pushing interest rates down. The resulting increase in the premium on public debt creates fiscal space: the government can service its debt more cheaply and sustain lower taxes while gradually moving the debt-to-GDP ratio towards the long-run target.[13]

5 Extensions

This section asks whether the properties of the optimal mix depend on the assumptions of the baseline analysis. I first solve the problem that ignores transitional dynamics, which makes it possible to use more standard welfare criteria (Section 5.1). I then allow for tax schedules

[13] Appendix E examines the sensitivity of these results to the approximation choices and discusses the restriction to smooth policy paths.

with lump-sum transfers (Section 5.2), introduce capital and taxes on savings (Section 5.3), and open the economy to foreign demand for safe assets (Section 5.4). In every case, planners with stronger preferences for redistribution choose more progressive tax systems and lower levels of public debt. The way the optimal mix responds to each change is itself informative about the mechanism: the negative correlation becomes *stronger* with capital, as redistribution now also operates through wages, and *weaker* in the open economy variant, where the interest rate channel is less prominent.

5.1 Optimal steady state analysis

The methods for solving the class of optimal policy problems in this paper have been developed fairly recently (**?**; **?**; **?**; **?**). To make progress, **?** focused on stationary outcomes, computing the level of debt that maximizes welfare in steady state. This remains a useful exercise: it is computationally tractable, it permits the study of richer models, and it can be thought of as the limit of the dynamic problem when the social discount factor approaches one.

Given a social welfare function $\mathscr{W}$, the optimal steady state problem (OSS) is to choose a time-invariant CRP tax code $\{\tau, p\}$ and the steady state level of public debt B in order to maximize

$$\max_{\{r,B,\tau,p\}} \mathscr{W}(r,\tau,p) \quad \text{s.t} \quad \begin{cases} \mathscr{A}^{ss}(r,\tau,p) = B, \\ G + rB = \mathscr{T}^{ss}(r,\tau,p) \end{cases} . \tag{11}$$

Here, $\mathscr{A}^{ss}(r,\tau,p)$ and $\mathscr{T}^{ss}(r,\tau,p)$ denote aggregate asset holdings and aggregate tax revenues in steady state. The two constraints ensure that the government budget is balanced and that the asset market clears.

I first use a social welfare function analogous to the one in the fully dynamic problem. That is, I assume the planner uses weights $\omega(a) \propto e^{-\alpha_a a}$ to aggregate instantaneous utilities. As before, I focus on how the optimal mix changes as the planner's aversion to wealth inequality α_a varies. The optimality conditions and the details behind the computations are in Appendix C.

Figure 5a traces out the optimal mix of debt and progressivity for the OSS problem (dashed black line) and compares it to the frontier of the fully dynamic problem (solid blue line). The

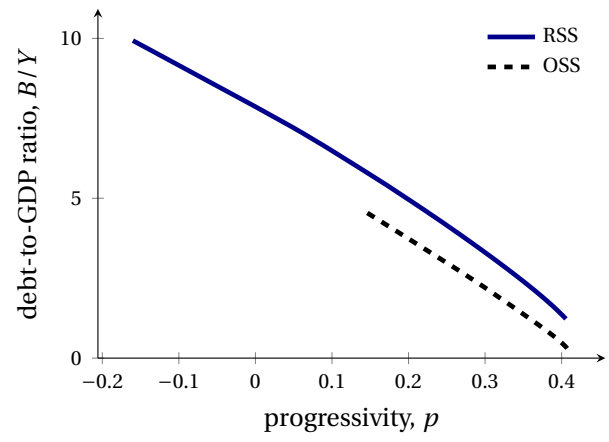
qualitative properties of the optimal mix are unchanged: planners with stronger preferences for redistribution implement more progressive tax systems and lower levels of debt. This also shows that the main results are robust to the choice of social discount factor. The reasons for the negative correlation between progressivity and indebtedness are the same as in the dynamic problem, as a similar sufficient statistics formula applies (see Appendix C.4).

Transitional dynamics only matter for the quantitative properties of the optimal mix. The OSS frontier shifts down and to the right: the optimal long-run mix in the OSS problem features relatively lower levels of public debt and higher progressivity.

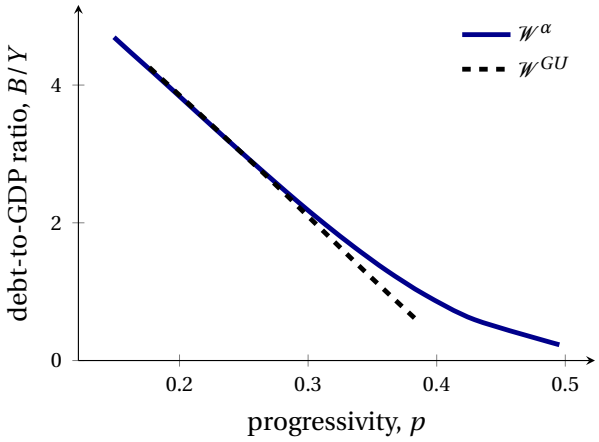
To see this more clearly, the bottom two panels of Figure 5 plot how each instrument varies with the planner's aversion to wealth inequality. For given α_a, the planner that ignores transitions favors relatively lower levels of public debt (bottom left panel). This is consistent with the results in **?**, who clarify how this concept of long-run optimality overestimates the costs of debt issuance and thus underestimates the optimal long-run level of public debt. The intuition for this result is that the planner that ignores transitions does not take into account the benefit of issuing debt along the transition, which helps reduce reliance on distortionary taxation. The bottom right panel shows that it also overestimates optimal long-run progressivity, for a similar reason: the planner does not internalize the distortionary effects of taxation along the transition.

The OSS problem also allows me to check that the results are not driven by my choice of SWF. I consider two alternatives, formally defined in Appendix C.5. The first is a generalized utilitarian criterion that aggregates *lifetime* utilities using weights that depend on asset holdings. Relative to a generational planner, this planner only cares about the inequality of the initial generation. The second, following **?** and **?**, aggregates consumption certainty-equivalents and separates society's aversion to inequality from households' aversion to intertemporal fluctuations. It nests a pure efficiency criterion, a utilitarian benchmark, and a Rawlsian planner who only cares about the poorest agents. The results are displayed in Figure 5b: whenever social preferences are such that the planner has a taste for redistribution, the optimal long-run mix features lower debt and higher progressivity. The negative correlation between progressivity and indebtedness is thus not driven by the details behind the SWF.

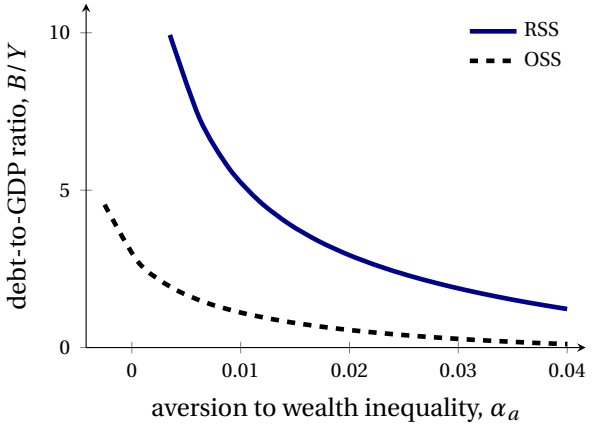
Figure 5: The optimal long-run mix ignoring transitions and across SWFs



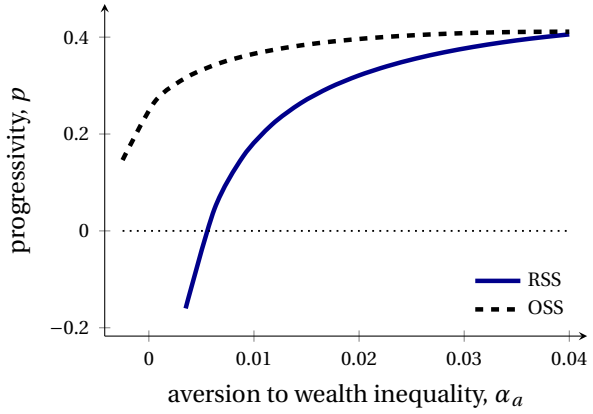
(a) Optimal long-run mix ignoring transitions



(b) Optimal long-run mix across SWFs



(c) Optimal B/Y with and without transitions



(d) Optimal p with and without transitions

5.2 Alternative labor income tax schedules

I also study the properties of the optimal mix with alternative labor income tax schedules. First, I consider a simpler tax system that still captures a form of progressive taxation: linear taxes with a lump-sum intercept. Specifically, I assume that taxes are now given by $T(y) = (1-\tau)y - T_0$, with $T_0 \geq 0$. For simplicity, I focus on the problem where the planner ignores transitions.

The results for this tax system are displayed in Figure 6 (solid black line). The main insight goes through: planners with stronger preferences for redistribution favor lower levels of debt. The reason is the interest rate channel that I have emphasized throughout the paper. An inequality-averse planner naturally favors lump-sum transfers because they lower average tax

rates at the bottom of the income distribution. But transfers, by acting as a form of social insurance against idiosyncratic risk, reduce the need for self-insurance. This lowers the aggregate demand for safe assets and puts upward pressure on interest rates, thereby making government borrowing more expensive.

Figure 6: Optimal long-run mix with transfers

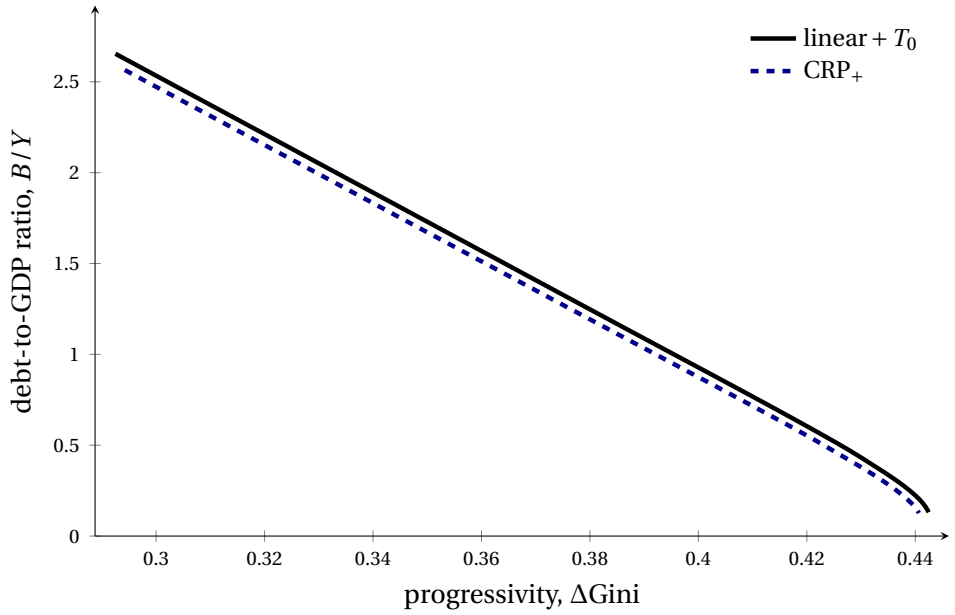


Figure 6 also presents the results for a three-parameter version of (1):

$$T(y) = y - \tau y^{1-p} - T_0,$$

where, once again, $T_0 \geq 0$. This CRP$_+$ tax system introduces a negative intercept into the tax system analyzed so far and gives the planner more flexibility. In particular, it allows the planner to disentangle average tax rates from marginal tax rates. A number of recent papers have argued that this improves the empirical fit to the overall tax and transfer system in the United States.[14]

As above, inequality-averse planners rely on lump-sum transfers to separate average tax rates from marginal tax rates and redistribute towards the poor. They continue to favor lower levels of debt, but unlike the baseline, the optimal tax system now features increasing average tax rates and *decreasing* marginal tax rates ($p < 0$). The intuition is that this combination

[14]The two-parameter tax function tends to overestimate taxes paid at the top and underestimate transfers at the bottom. See **?** and **?** for more details.

achieves redistribution while preserving efficiency (**?**). With transfers, p is no longer a sufficient statistic for the progressivity of the tax system. A more global measure is needed, such as the change in the Gini coefficient between the before-tax and after-tax income distributions. Once progressivity is measured this way, the negative correlation between progressivity and indebtedness remains (notice the x-axis in Figure 6).

5.3 Multiple safe assets and taxes on savings

In the baseline model, the only supply of bonds outside the household sector comes from the government. This section considers an extension with a more general production technology $F(K,L)$ and allows firms to issue claims to capital. This introduces an alternative asset that households can use to smooth their consumption and brings the model closer to the one studied by **?**. The details behind the calibration of this model and a complete definition of the equilibrium are relegated to Appendix D.1.

Given the absence of aggregate risk, capital and government bonds are perfect substitutes. This means that in equilibrium the interest rate on both assets must be the same. Firm optimality implies that before-tax factor prices w_t and r_t must satisfy

$$w_t = F_L(K_{t-1}, L_t) \quad \text{and} \quad r_t = F_K(K_{t-1}, L_t) - \delta.$$

The key difference with respect to the baseline model is that the asset market clearing condition becomes

$$B_t + K_t = \int \boldsymbol{a}_t(x) dD_t(x).$$

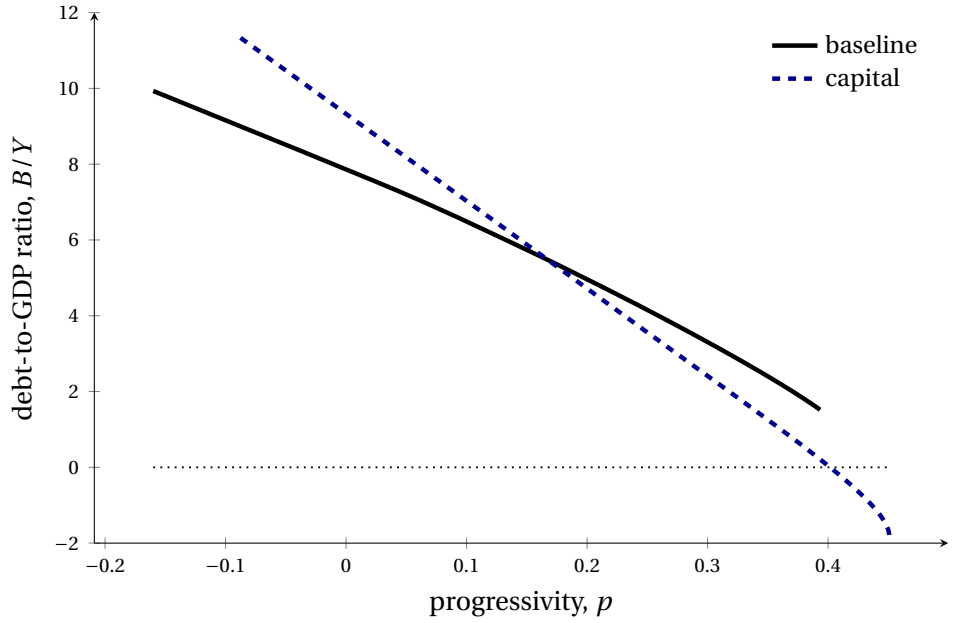
This implies that the supply of safe assets is not perfectly inelastic. The interest-rate elasticity of the supply of safe assets will of course depend on the curvature of the production function.

Here, I work with a standard Cobb-Douglas production function $F(K,L) = K^{\alpha}L^{1-\alpha}$, which generates a fairly inelastic supply of capital. This has implications for the interest rate channel discussed in Section 3. Specifically, the response of the equilibrium interest rate to changes in the progressivity of the tax system becomes less pronounced. There is an additional margin of adjustment and thus the change in interest rates that is required to clear the asset market is smaller than in the baseline (see Figure 17 in Appendix D.1).

The dynamic optimal policy problem for this economy is similar to the one in Section 4 except that I allow the planner to tax the return on savings at rate τ_k.[15] In Appendix D.2, I derive the optimality conditions for this problem and show that the planner chooses τ_k in order to implement the modified golden rule ($F_K = \delta + \beta^{-1} - 1$). This result is already in **?** but I verify that it goes through when the planner has a taste for redistribution. As **?** discuss, the fact that distributional concerns do not interfere with the efficient level of investment is reminiscent of the production efficiency result in **?**. There is no need to implement an inefficiently high level of capital to help agents self-insure against risk because of the availability of public debt.

Once the capital-labor ratio has been determined, the problem essentially reduces to a reparameterized version of the baseline model. This helps understand why the qualitative properties of the optimal long-run mix, displayed in Figure 7, are similar to the baseline.

Figure 7: Optimal long-run mix with capital



Interestingly, the negative correlation between progressivity and debt becomes *stronger*: the dashed blue line is steeper than the solid black line. This is despite the fact that the interest rate channel is less powerful in the model with capital. The reason is an effect that operates through wages, which is absent in the baseline model because wages are fixed. An increase in the progressivity of the tax system increases r but lowers w. This scales down the stochastic

[15] A linear tax on savings makes no difference in the baseline model.

component of household income, so the value of providing public liquidity falls by more than in the baseline. Relatedly, lowering B frees savings for capital and raises the wage. Since the poor rely relatively more on labor income, planners that care about redistribution value this and lower debt by more.

The quantitative differences arise because the supply of capital reduces the reliance on government debt to satisfy the demand for safe assets. In addition, debt now crowds out capital, an additional cost that calls for lower levels of debt in the long run.

5.4 Open economy: a direct test of the mechanism

I now extend the baseline model to allow for foreign demand for safe assets. This provides a direct test of the mechanism: if the interest rate channel drives the negative correlation between progressivity and indebtedness, weakening the channel should weaken the correlation. Following **?**, I introduce debt-elastic interest rates. This means the domestic interest rate is now given by

$$r = r^{\text{ROW}} + \psi_2(e^{D_f} - 1), \tag{12}$$

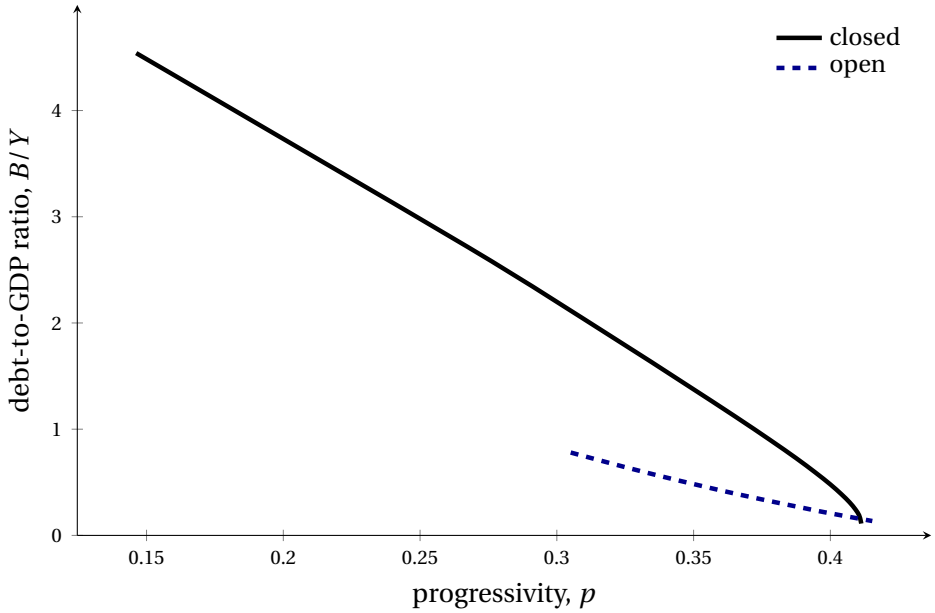
where r^{ROW} is the world interest rate, D_f is net foreign debt, and $\psi_2 \geq 0$ governs the sensitivity of domestic interest rates to net foreign debt. The parameter ψ_2 controls the strength of the interest rate channel: when $\psi_2 \to 0$, the economy approaches the small open economy limit with $r = r^{\text{ROW}}$, effectively shutting down the interest rate channel. As ψ_2 increases, the model behaves more like the closed economy baseline.

The asset market clearing condition becomes $A + D_f = B$, where $A = \int a_i di$ is domestic private asset demand and B is government debt. The government's budget constraint remains unchanged.

Figure 8 shows the optimal mix of debt and progressivity across different values of inequality aversion for both the closed economy (solid blue line) and a partially open economy (dashed black line). The negative relationship between progressivity and debt becomes less pronounced.

This supports the idea that the interest rate channel is a key driver of the properties of the optimal mix. With foreign demand for safe assets, the link between domestic progressivity and equilibrium interest rates weakens. The negative relationship survives but becomes weaker,

Figure 8: Optimal long-run mix: closed vs. open economy



Note: The figure shows the optimal mix of debt and progressivity for different values of inequality aversion. The solid blue line represents the closed economy benchmark, while the dashed black line represents an open economy with debt-elastic interest rates ($\psi_2 = 0.00005, r^{\text{ROW}} = 0$).

exactly as the mechanism predicts.

6 Inverting the optimum

Section 4 read the negative association among advanced economies as variation in preferences. This section pushes that reading further. I first ask how a utilitarian planner evaluates the US mix of debt and progressivity. I then recover, for each advanced economy, the preference for redistribution that makes its observed mix optimal. I do not claim that governments set fiscal policy by solving Ramsey problems. Instead, the exercise treats the model as a measuring device. Applied country by country, it turns observed policy into a ranking of preferences for redistribution. The ranking is stable across specifications and lines up with independent evidence.

6.1 Utilitarian planners & US fiscal policy

From the perspective of a utilitarian planner, the US issues far too little debt. The average value of public debt in the US between 1995 and 2007 was around 61.5% of GDP (**?**), well below 290%, the optimal ratio in the steady-state problem with a utilitarian SWF analyzed in Section 5.[16] If anything, this understates the gap: the steady-state problem overstates the costs of issuing debt and thus understates its optimal long-run level (Section 5.1).

The verdict on progressivity depends on whether the level of debt is part of the problem. Taking the observed level of debt as given, the utilitarian planner would make the tax system substantially more progressive: the restricted optimum features a progressivity of 0.375, well above the 0.181 estimated by **?** using PSID data from 2000 to 2006, and in line with the literature that abstracts from public debt (**?**). Once debt is chosen freely, this is no longer the case. The planners closest to the utilitarian benchmark choose regressive tax schedules (Section 4), and in the steady-state problem, where the utilitarian optimum exists, optimal progressivity falls to 0.25, close to the estimated value. Ignoring the debt margin thus overestimates the optimal progressivity of the tax system. The reason is the interest rate channel of Section 3: it pushes the planner toward less progressivity once the level of debt is part of the problem.

These large differences are not driven by the absence of capital in the baseline model. In the model with multiple safe assets, the picture is the same. First, if the planner is not allowed to tax capital, the optimal level of debt is closer to what we see in the data, but the gap between optimal and estimated progressivity *increases*. Second, in the model with capital taxes analyzed in Section 5.3, optimal progressivity remains well below the estimated value and the optimal level of debt remains well above the data. Allowing for more flexible forms of labor income taxation (i.e. lump-sum transfers as in Section 5.2) does not change the answer.

6.2 The implied ranking of preferences for redistribution

A natural question is what preferences would rationalize the policies we observe. Varying the taste for redistribution alone is not enough: it traces out the one-dimensional frontier of Sec-

[16] The baseline calibration in Section 2 targets 140% instead, because debt is the only safe asset in that model and the notion is broader.

tion 4, and the aversion that matches a country's level of debt generally implies the wrong progressivity. I therefore let income risk vary as well. For each country, I choose the pair (α_a, σ) such that the Ramsey steady state reproduces the observed progressivity and debt-to-GDP ratio, holding every other primitive at the baseline calibration. Because preferences and risk move the optimal mix in opposite directions, the two observed policies pin down the two parameters. The exercise treats each country as the same baseline economy with a different planner and a different amount of income risk.

The progressivity targets come from **?**, who estimate the tax function (1) on microdata from the Luxembourg Income Study. The estimates capture taxes and transfers together, which is the concept in the model. The debt targets are general government gross debt from the IMF, averaged between 1995 and 2007. I restrict attention to economies that were already advanced in 1996, the comparison group of Section 4.4.

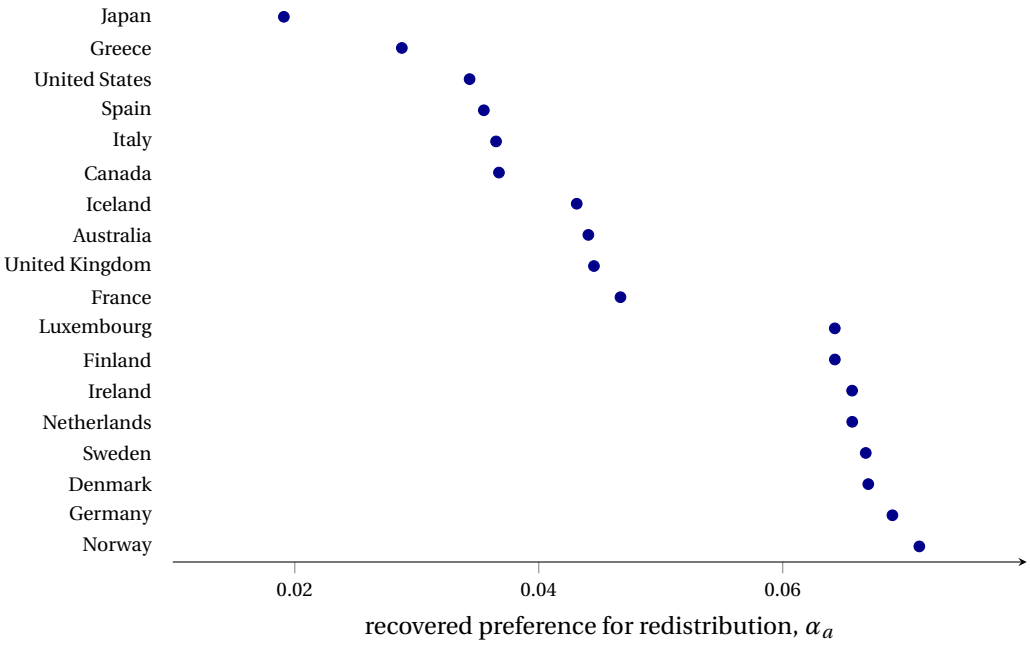
The implied α_a ranges from 0.019 in Japan to 0.071 in Norway (Figure 9). The US is near the bottom at 0.034, while Denmark is near the top at 0.067. Allowing households to hold capital leaves the ranking almost unchanged.

The US and Denmark land at opposite ends of the ranking. More generally, Japan and the liberal market economies sit toward the bottom, Germany and the Nordic economies toward the top. This is qualitatively consistent with the cross-country inversion of **?** and with survey and experimental evidence on redistributive preferences (**?**; **?**).

Greece and Italy are the exceptions: the model sees high debt and relatively low progressivity and recovers a low α_a. In both countries, political economy considerations arguably weigh more heavily on fiscal policy. The low values there likely reflect forces outside the model rather than a weak preference for redistribution.

Even if no government solves a Ramsey problem when setting fiscal policy, the model provides a consistent way to compare fiscal policy across countries. Through the lens of the model, it is as if the advanced economies differed mainly in their preferences for redistribution, and the ranking says by how much.

Figure 9: Model-implied preferences for redistribution across advanced economies



Note: Each point reports the recovered value of α_a. The policy targets use the progressivity estimates of **?** and general government gross debt.

7 Conclusion

This paper studies the optimal mix of public debt and redistributive taxation in models with heterogeneous agents and incomplete markets. I find that planners with stronger preferences for redistribution implement more progressive tax systems and, perhaps surprisingly, favor lower levels of public debt. This is due to an interest rate channel: progressive income taxes reduce the need to self-insure against idiosyncratic risk, lowering the demand for safe assets and increasing the fiscal cost of public debt. This result is robust to the presence of multiple safe assets and, in the optimal steady-state exercises, to alternative social welfare functions and more flexible tax schedules.

Along the optimal transition, the interest rate channel continues to play a central role. The planner starts by flattening the tax schedule, reducing the insurance provided by the tax system. This pushes interest rates down and creates fiscal space, allowing the government to sustain lower taxes and higher consumption while gradually moving the debt-to-GDP ratio towards the

long-run target. Ignoring transitional dynamics underestimates the optimal level of public debt and overestimates the optimal degree of progressivity.

The analysis also suggests one way of reading the cross-country evidence. Among advanced economies, the relationship between debt and progressivity is negative, the sign that variation in preferences for redistribution would generate. Inverting the optimum delivers a ranking of preferences for redistribution that lines up with independent evidence.

On the methodological side, I extend the sequence-space approach of **?** to compute optimal transitions via policy gradient methods. Allowing for aversion to wealth inequality helps overcome the difficulties that arise when searching for a Ramsey steady state in these models.

Finally, the analysis abstracts from political economy considerations, which are likely to shape the planner's ability to commit to a long-run policy mix. This is an important direction for future work. Nevertheless, the results make clear that the level of public debt and the progressivity of the tax system should not be set independently: a government's stance on redistribution directly affects its capacity to borrow.

A Appendix to interest rate channel

This appendix decomposes the effects of a progressive tax reform on individual lifetime utilities across the entire state space. The second part of Lemma 1 shows that changes in the progressivity of the tax system affect individual lifetime utilities through partial equilibrium (direct) and general equilibrium (indirect) channels. To see this more clearly, iterate (5) forward to write the response of individual outcomes as

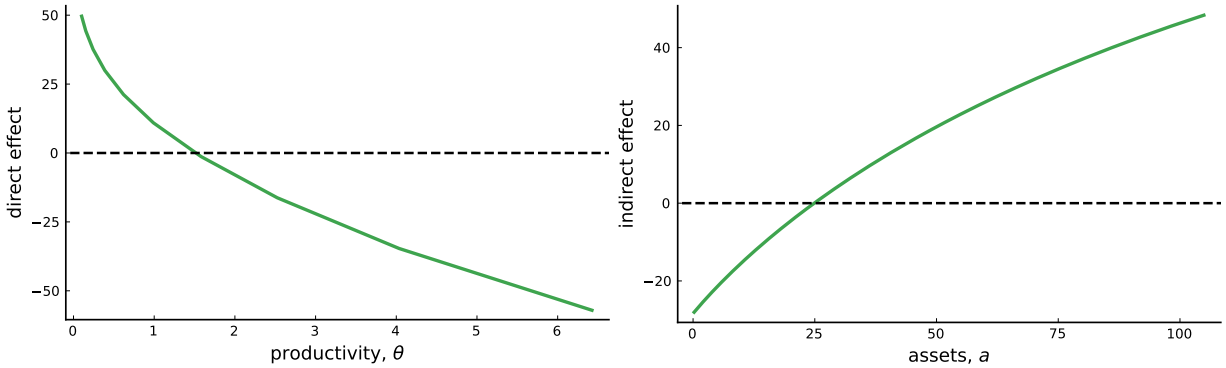
$$dV(x) = \sum_{s=0}^{\infty} \beta^s \mathbb{E}\left[u'(c_s) \left(\underbrace{y_s^{1-p} d\boldsymbol{\tau} + a_s \, d\boldsymbol{r}}_{\text{indirect effect in } s} - \underbrace{z_s \log y_s \, dp}_{\text{direct effect in } s} \right) \mid x_0 = x \right].$$

The second term inside the parenthesis is the expected direct effect s periods after the reform for an agent with initial state equal to x. This comes about because the change in p affects the slope of the tax schedule. By doing so, it lowers taxes paid at the bottom of the income distribution while increasing the taxes paid at the top. Figure 10a shows that this partial equilibrium effect tends to favor the agents that are income-poor (i.e. those with low productivity) at the time of the reform.

The first term inside the parenthesis is the expected indirect effect s periods after the reform, which shows up due to the general equilibrium nature of the exercise: changes in progressivity trigger movements in the real interest rate and the level of taxes in order to ensure asset markets clear and the government's constraint holds. As Figure 10b shows, these general equilibrium effects tend to favor the agents that are asset-rich at the time of the reform. These effects make it possible for those at the top of the wealth distribution to *benefit* from progressive tax reforms. This result is driven mainly by the positive interest rate response described in Section 3.

By combining the logic behind the direct and indirect effects, the state space can be divided into (a) regions where agents unambiguously win or lose from the reform, and (b) regions where the outcome is ambiguous. First, those who are both income-poor and asset-rich will always favor progressive tax reforms: they benefit from lower taxes at the bottom of the income distribution and the higher interest rates induced by the reform. Second, those who are both income-rich and asset-poor will always be worse off: they pay higher taxes and do not have enough assets to benefit from the increase in interest rates. For the agents that are poor or rich along both dimensions, the ultimate effect is ambiguous and hinges on whether the direct or

Figure 10: Direct and indirect effects across the state space



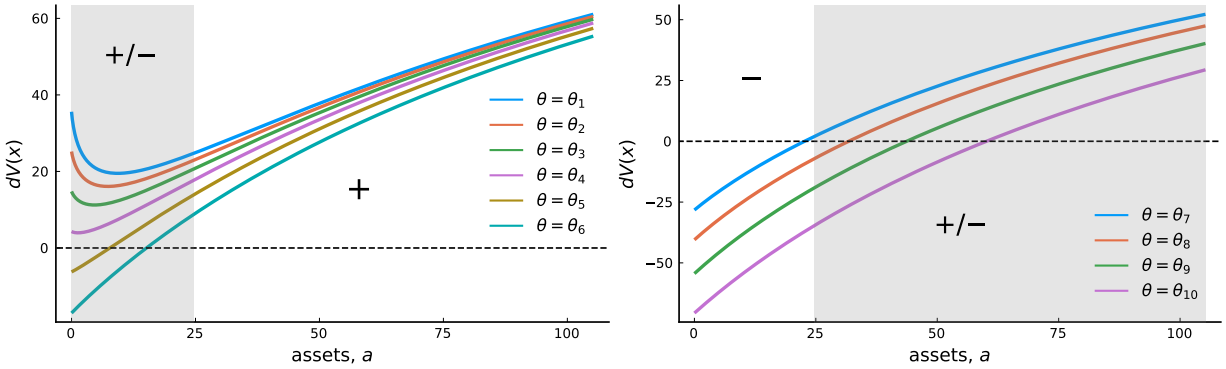
(a) Direct effect along the θ dimension

(b) Indirect effect along the a dimension

Note: The direct effect is integrated over assets using the stationary marginal distribution of assets, while the indirect effect is integrated over productivity using the stationary marginal distribution of productivity.

the indirect effect dominates, which is a quantitative question.

Figure 11: Individual responses across the state space



(a) Total effect at the bottom of θ distribution

(b) Total effect at the top of θ distribution

These different regions can be seen in Figure 11, where I plot the total change in the value function along the asset dimension, for different levels of productivity. The panel on the left shows the change in lifetime utilities for the agents at the bottom of the productivity distribution, whereas the panel on the right shows it for those at the top. The figures are constructed using the thresholds for assets and productivity where the indirect and direct effect switch signs.

Productivity levels that are below the value of θ at which the direct effect becomes negative are included in the left panel, whereas the productivity levels that are above this threshold are in the right panel. The frontier between the grey region, where the total effect is ambiguous, and the white region, where there is no ambiguity, corresponds to the level of assets where the indirect effect switches sign.

The asset-rich who are also income-poor at the time of the reform always benefit from progressive income taxes: the overall change in the value function is positive. Moreover, in this calibration, the interest rate response is so dominant that it allows the asset-rich to benefit from the reform, regardless of whether they are at the bottom or at the top of the productivity distribution. On the other hand, most of the agents with low wealth end up losing from the reform, except for those at the bottom of the productivity distribution. For the asset-poor, the indirect effect is mostly driven by the effect of the reform on the level of taxes, which is negative. Among these agents, only the low-productivity types are better off because they alone benefit from the direct effects of the reform.

B Appendix to Ramsey problem

This appendix provides additional details on the Ramsey problem and its solution. I start by deriving necessary conditions for the Ramsey steady state by following the sequence-space approach recently introduced by **?**. I then describe the algorithm to compute the optimal long-run mix of debt and progressivity. Finally, I discuss alternative formulations of the RSS problem and derive the sufficient-statistic representation of the optimal long-run value of public debt in Proposition 1.

B.1 Optimality conditions for RSS problem

Recall that the dynamic Ramsey problem is

$$\max_{\{r_t,B_t,p_t,\tau_t\}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) \quad \text{s.t} \quad \begin{cases} \mathscr{A}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) = B_t, \\ G + (1+r_{t-1})B_{t-1} = B_t + \mathscr{T}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) \end{cases}, \forall\, t.$$

Given *sequences* for debt $\{B_t\}$ and progressivity $\{p_t\}$, one can take into account the period-by-period constraints by solving for *sequence space functions* $\boldsymbol{r}_t(\{p_s\},\{B_s\})$ and $\boldsymbol{\tau}_t(\{p_s\},\{B_s\})$. Then, the problem becomes

$$\max_{\{p_t,B_t\}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{\boldsymbol{r}_s\left(\{p_u\},\{B_u\}\right)\},\{\boldsymbol{\tau}_s\left(\{p_u\},\{B_u\}\right)\},\{p_s\}\right).$$

Given an optimal plan $\{p_t^*, B_t^*\}$, a necessary condition for optimality is that any perturbation dp_u and dB_u shouldn't affect welfare. This yields the following pair of optimality conditions for any $u = 0, 1, \ldots$:

$$\sum_{t=0}^{\infty} \beta^{t-u} \sum_{s=0}^{\infty} \frac{\partial \mathscr{U}_t}{\partial r_s} \frac{\partial \boldsymbol{r}_s}{\partial p_u} + \sum_{t=0}^{\infty} \beta^{t-u} \sum_{s=0}^{\infty} \frac{\partial \mathscr{U}_t}{\partial \tau_s} \frac{\partial \boldsymbol{\tau}_s}{\partial p_u} + \sum_{t=0}^{\infty} \beta^{t-u} \frac{\partial \mathscr{U}_t}{\partial p_u} = 0 \tag{13}$$

$$\sum_{t=0}^{\infty} \beta^{t-u} \sum_{s=0}^{\infty} \frac{\partial \mathscr{U}_t}{\partial r_s} \frac{\partial \boldsymbol{r}_s}{\partial B_u} + \sum_{t=0}^{\infty} \beta^{t-u} \sum_{s=0}^{\infty} \frac{\partial \mathscr{U}_t}{\partial \tau_s} \frac{\partial \boldsymbol{\tau}_s}{\partial B_u} = 0 \tag{14}$$

To simplify (13) and (14) we take $u \to \infty$. Now, for any sequence space function $F_t(\{X_s\})$, define the discounted sum $\mathscr{E}_X^F$ of the long-run response,

$$\mathscr{E}_X^F \equiv \lim_{u\to\infty} \sum_{t=0}^{\infty} \beta^{t-u} \frac{\partial F_t}{\partial X_u}.$$

Now, combining the *quasi-Toeplitz* property of Jacobians in stationary models[17] with the convolution theorem, we can write the composition of Jacobians as the product of discounted sums. That is,

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\sum_{s=0}^{\infty}\frac{\partial \mathcal{U}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial p_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathcal{U}_t}{\partial r_u}\cdot\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \boldsymbol{r}_t}{\partial p_u}=\mathcal{E}_r^{\mathcal{U}}\cdot\mathcal{E}_p^{\boldsymbol{r}} \tag{15}$$

To see why this is the case, observe that by defining $m=t-u$ and $n=s-u$,

$$\sum_{t=0}^{\infty}\beta^{t-u}\sum_{s=0}^{\infty}\frac{\partial \mathcal{U}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial p_u}=\sum_{m=-u}^{\infty}\beta^{m}\sum_{n=-u}^{\infty}\frac{\partial \mathcal{U}_{m+u}}{\partial r_{n+u}}\frac{\partial \boldsymbol{r}_{n+u}}{\partial p_u}$$

The quasi-Toeplitz property of the Jacobians implies that as $u\to\infty$,

$$\sum_{m=-u}^{\infty}\beta^{m}\sum_{n=-u}^{\infty}\frac{\partial \mathcal{U}_{m+u}}{\partial r_{n+u}}\frac{\partial \boldsymbol{r}_{n+u}}{\partial p_u}\to\sum_{m=-\infty}^{\infty}\beta^{m}\sum_{n=-\infty}^{\infty}f_{m-n}g_n,$$

where $\boldsymbol{f}=\{f_t\}$ and $\boldsymbol{g}=\{g_t\}$ are the symbol vectors of $\frac{\partial \mathcal{U}_t}{\partial r_s}$ and $\frac{\partial \boldsymbol{r}_t}{\partial p_s}$, respectively. The term inside the summation is the convolution of the functions $\boldsymbol{f}$ and $\boldsymbol{g}$. Since quasi-Toeplitz Jacobians decay at least exponentially away from the diagonal, $\{f_t\}$ and $\{g_t\}$ are summable. The convolution theorem then implies that

$$\sum_{m=-\infty}^{\infty}\beta^{m}\sum_{n=-\infty}^{\infty}f_{m-n}g_n=\sum_{k=-\infty}^{\infty}\beta^{k}f_k\sum_{k=-\infty}^{\infty}\beta^{k}g_k$$

Now,

$$\sum_{k=-\infty}^{\infty}\beta^{k}f_k=\sum_{k=-\infty}^{\infty}\beta^{k}\lim_{u\to\infty}\frac{\partial \mathcal{U}_{k+u}}{\partial r_u}=\lim_{u\to\infty}\sum_{k=-u}^{\infty}\beta^{k}\frac{\partial \mathcal{U}_{k+u}}{\partial r_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathcal{U}_t}{\partial r_u}=\mathcal{E}_r^{\mathcal{U}}\quad,$$

where the second equality follows from the dominated convergence theorem and the third equality by substituting $t=k+u$. Analogously,

$$\sum_{k=-\infty}^{\infty}\beta^{k}g_k=\sum_{k=-\infty}^{\infty}\beta^{k}\lim_{u\to\infty}\frac{\partial \boldsymbol{r}_{k+u}}{\partial p_u}=\lim_{u\to\infty}\sum_{k=-u}^{\infty}\beta^{k}\frac{\partial \boldsymbol{r}_{k+u}}{\partial p_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \boldsymbol{r}_t}{\partial p_u}=\mathcal{E}_p^{\boldsymbol{r}}.$$

This establishes (15). Following similar steps, one can show that

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\sum_{s=0}^{\infty}\frac{\partial \mathcal{U}_t}{\partial \tau_s}\frac{\partial \boldsymbol{\tau}_s}{\partial p_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathcal{U}_t}{\partial \tau_u}\cdot\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \boldsymbol{\tau}_t}{\partial p_u}=\mathcal{E}_\tau^{\mathcal{U}}\cdot\mathcal{E}_p^{\boldsymbol{\tau}}$$

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\sum_{s=0}^{\infty}\frac{\partial \mathcal{U}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial B_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathcal{U}_t}{\partial r_u}\cdot\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \boldsymbol{r}_t}{\partial B_u}=\mathcal{E}_r^{\mathcal{U}}\cdot\mathcal{E}_B^{\boldsymbol{r}}$$

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\sum_{s=0}^{\infty}\frac{\partial \mathcal{U}_t}{\partial \tau_s}\frac{\partial \boldsymbol{\tau}_s}{\partial B_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathcal{U}_t}{\partial \tau_u}\cdot\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \boldsymbol{\tau}_t}{\partial B_u}=\mathcal{E}_\tau^{\mathcal{U}}\cdot\mathcal{E}_B^{\boldsymbol{\tau}}.$$

[17] An infinite matrix $\boldsymbol{J}=\{J_{t,s}\}\in\mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ is a quasi-Toeplitz matrix with symbol vector $\boldsymbol{j}=\{j_t\}\in\mathbb{R}^{\mathbb{Z}}$ if i) for any $t,s\geq 0$ we have $\lim_{u\to\infty}J_{t+u,s+u}=j_{t-s}$ and ii) the entries of $J_{t,s}$ decay at least exponentially away from the diagonal. See **?** and **?** for details.

This yields two necessary conditions for the Ramsey steady state

$$\mathscr{E}_r^{\mathscr{U}} \cdot \mathscr{E}_p^{\boldsymbol{r}} + \mathscr{E}_\tau^{\mathscr{U}} \cdot \mathscr{E}_p^{\boldsymbol{\tau}} + \mathscr{E}_p^{\mathscr{U}} = 0 \tag{16}$$

$$\mathscr{E}_r^{\mathscr{U}} \cdot \mathscr{E}_B^{\boldsymbol{r}} + \mathscr{E}_\tau^{\mathscr{U}} \cdot \mathscr{E}_B^{\boldsymbol{\tau}} = 0 \tag{17}$$

The discounted sum of the general equilibrium long-run responses $\mathscr{E}_p^{\boldsymbol{r}}$, $\mathscr{E}_p^{\boldsymbol{\tau}}$, $\mathscr{E}_B^{\boldsymbol{r}}$, and $\mathscr{E}_B^{\boldsymbol{\tau}}$ can then be computed from the system of equations derived below.

System for GE responses

I start by deriving a system of equations for $\mathscr{E}_p^{\boldsymbol{r}}$ and $\mathscr{E}_p^{\boldsymbol{\tau}}$. Perturbing the asset market-clearing condition at time t by dp_u,

$$\sum_{s=0}^{\infty} \frac{\partial \mathscr{A}_t}{\partial r_s} \frac{\partial \boldsymbol{r}_s}{\partial p_u} + \sum_{s=0}^{\infty} \frac{\partial \mathscr{A}_t}{\partial \tau_s} \frac{\partial \boldsymbol{\tau}_s}{\partial p_u} + \frac{\partial \mathscr{A}_t}{\partial p_u} = 0, \ \forall t.$$

Multiply condition at time t by β^{t-u} and sum across t to get

$$\sum_{t=0}^{\infty} \beta^{t-u} \sum_{s=0}^{\infty} \frac{\partial \mathscr{A}_t}{\partial r_s} \frac{\partial \boldsymbol{r}_s}{\partial p_u} + \sum_{t=0}^{\infty} \beta^{t-u} \sum_{s=0}^{\infty} \frac{\partial \mathscr{A}_t}{\partial \tau_s} \frac{\partial \boldsymbol{\tau}_s}{\partial p_u} + \sum_{t=0}^{\infty} \beta^{t-u} \frac{\partial \mathscr{A}_t}{\partial p_u} = 0.$$

Letting $u \to \infty$ and using the same results we relied on when simplifying the FOCs, this becomes

$$\mathscr{E}_r^{\mathscr{A}} \cdot \mathscr{E}_p^{\boldsymbol{r}} + \mathscr{E}_\tau^{\mathscr{A}} \cdot \mathscr{E}_p^{\boldsymbol{\tau}} + \mathscr{E}_p^{\mathscr{A}} = 0.$$

Following the same steps with the government's budget constraint yields

$$\left(\mathscr{E}_r^{\mathscr{T}} - B\right) \mathscr{E}_p^{\boldsymbol{r}} + \mathscr{E}_\tau^{\mathscr{T}} \cdot \mathscr{E}_p^{\boldsymbol{\tau}} + \mathscr{E}_p^{\mathscr{T}} = 0.$$

Thus, we can solve for $\mathscr{E}_p^{\boldsymbol{r}}$ and $\mathscr{E}_p^{\boldsymbol{\tau}}$ from

$$\begin{bmatrix} \mathscr{E}_r^{\mathscr{A}} & \mathscr{E}_\tau^{\mathscr{A}} \\ \mathscr{E}_r^{\mathscr{T}} - B & \mathscr{E}_\tau^{\mathscr{T}} \end{bmatrix} \begin{bmatrix} \mathscr{E}_p^{\boldsymbol{r}} \\ \mathscr{E}_p^{\boldsymbol{\tau}} \end{bmatrix} = \begin{bmatrix} -\mathscr{E}_p^{\mathscr{A}} \\ -\mathscr{E}_p^{\mathscr{T}} \end{bmatrix}. \tag{18}$$

Analogous derivations for a perturbation dB_u lead to a system for $\mathscr{E}_B^{\boldsymbol{r}}$ and $\mathscr{E}_B^{\boldsymbol{\tau}}$

$$\begin{bmatrix} \mathscr{E}_r^{\mathscr{A}} & \mathscr{E}_\tau^{\mathscr{A}} \\ \mathscr{E}_r^{\mathscr{T}} - B & \mathscr{E}_\tau^{\mathscr{T}} \end{bmatrix} \begin{bmatrix} \mathscr{E}_B^{\boldsymbol{r}} \\ \mathscr{E}_B^{\boldsymbol{\tau}} \end{bmatrix} = \begin{bmatrix} 1 \\ \beta(1+r) - 1 \end{bmatrix}. \tag{19}$$

B.2 Computing the RSS

The algorithm to solve for the RSS relies on (16) and (17). To operationalize these equations, it is necessary to compute the discounted sum of the long-run response of aggregate outcomes $\mathscr{Y} \in \{\mathscr{U}, \mathscr{A}, \mathscr{T}\}$ to changes in interest rates and changes in the level and the progressivity of the tax system. In principle, this could be done by computing the Jacobians for e.g. aggregate welfare, assets, and taxes around the steady state implied by some candidate fiscal policy, and then taking the discounted sum of some far out column. However, this turns out to be too costly since we need a pretty large horizon to get convergence of the discounted sum. Given this, **?** propose the following procedure for the long-run responses. For concreteness, I focus on the discounted sum of the long-run response of aggregate asset demand to changes in interest rates $\mathscr{E}_r^{\mathscr{A}}$. The same procedure applies to the other responses.

- Iterate backward to find perturbation $d\boldsymbol{a}^s(\theta, a)$ to policy function when shock dr is $s = 0, 1, \ldots$ periods in the future. Then, sum across periods

$$d\boldsymbol{a}(\theta, a) = \sum_{s=0}^{\infty} \beta^{-s} d\boldsymbol{a}^s(\theta, a).$$

- Using $d\boldsymbol{a}(\theta, a)$ as the perturbation to asset policy function, iterate forward to find implied change in distribution D for $s = 1, 2, \ldots$ periods in the future. Take sum

$$dD(\theta, a) = \sum_{s=1}^{\infty} \beta^s dD^s(\theta, a).$$

- Finally, aggregate $dA = d\boldsymbol{a} \cdot D^{ss} + dD \cdot \boldsymbol{a}^{ss}$.[18] Then, $\mathscr{E}_r^{\mathscr{A}} = \frac{dA}{dr}$.

With this in hand, I solve for the RSS as follows:

1. Given a candidate level of public debt B and progressivity p, solve for the interest rate r and the level of taxes τ that ensure asset market clearing and government budget balance.

2. Compute the discounted sum of the long-run responses of aggregate assets $\mathscr{E}_r^{\mathscr{A}}, \mathscr{E}_\tau^{\mathscr{A}}, \mathscr{E}_p^{\mathscr{A}}$, aggregate utility $\mathscr{E}_r^{\mathscr{U}}, \mathscr{E}_\tau^{\mathscr{U}}, \mathscr{E}_p^{\mathscr{U}}$, and taxes $\mathscr{E}_r^{\mathscr{T}}, \mathscr{E}_\tau^{\mathscr{T}}, \mathscr{E}_p^{\mathscr{T}}$ using the procedure outlined above.

[18] Here, D^{ss} and $\boldsymbol{a}^{ss}$ are the distribution and policy function in the steady state implied by the candidate fiscal policy.

3. Solve for the discounted sum of the general equilibrium long-run responses using (18) and (19).

4. Check whether (16) and (17) are satisfied. If not, update r and p and repeat steps 1-4.

In practice, to reduce the complexity of general equilibrium, it is better to iterate on r and p, and then use asset market clearing to read off the implied level of public debt B. This means that in step (1) above, we only need to solve for the level of taxes τ that ensures government budget balance. Moreover, I use a two-step procedure to avoid multi-dimensional root finding algorithms. So in step (4), I first check whether (16) is satisfied. If this is not the case, I adjust p and repeat the process but keeping r fixed. Because this is a one-dimensional problem, the updating for p can be done via Brent's method. Once (16) is satisfied at the candidate level of r, I check whether (17) is satisfied. If this is not the case, I adjust r and repeat the process, resolving for p along the way. As explained in the main text, I verify that there is a unique interior solution to (17), fixing p. I also verify that there is a unique solution to (16). See Appendix B.6 for a figure.

B.3 Alternative formulation of the RSS problem

To simplify the derivations and side-step the complexity introduced by the presence of general equilibrium responses in (13) and (14), we can proceed as follows. By Walras' Law, problem (6) is equivalent to

$$\max_{\{r_t,B_t,p_t,\tau_t\}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) \quad \text{s.t} \quad \begin{cases} \mathscr{A}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) = B_t, \\ G + \mathscr{C}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) = \mathscr{Y}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) \end{cases} \forall t.$$

Now, given choices of $\{r_t,\tau_t,p_t\}$, the asset market clearing condition pins down $\{B_t\}$. After dropping all redundant choice variables and the associated constraints, the problem reduces to

$$\max_{\{r_t,p_t,\tau_t\}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) \quad \text{s.t} \quad G + \mathscr{C}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) = \mathscr{Y}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right).$$

The Lagrangian for this problem is

$$\max_{\{r_t,\tau_t,p_t\}} \sum_{t=0}^{\infty} \beta^t \left\{\mathscr{U}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) + \lambda_t^{GM}\{\mathscr{Y}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) - \mathscr{C}_t\left(\{r_s\},\{\tau_s\},\{p_s\}\right) - G\}\right\}$$

The first-order conditions are

$$\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial r_u}+\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\left(\frac{\partial \mathscr{Y}_t}{\partial r_u}-\frac{\partial \mathscr{C}_t}{\partial r_u}\right)=0,\quad \text{for } u=0,1,\ldots \tag{20}$$

$$\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial \tau_u}+\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\left(\frac{\partial \mathscr{Y}_t}{\partial \tau_u}-\frac{\partial \mathscr{C}_t}{\partial \tau_u}\right)=0,\quad \text{for } u=0,1,\ldots \tag{21}$$

$$\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial p_u}+\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\left(\frac{\partial \mathscr{Y}_t}{\partial p_u}-\frac{\partial \mathscr{C}_t}{\partial p_u}\right)=0,\quad \text{for } u=0,1,\ldots, \tag{22}$$

together with goods-market clearing period-by-period. To characterize the steady state, we take $u\to\infty$ and assume the multipliers converge: $\lambda_t^{GM}\to\lambda^{GM}>0$ as $t\to\infty$. The first term in (20) clearly goes to $\mathscr{E}_r^{\mathscr{U}}$. Now consider the second term, $\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\frac{\partial \mathscr{Y}_t}{\partial r_u}$. Since $\frac{\partial \mathscr{Y}_t}{\partial r_s}$ is a quasi-Toeplitz matrix with some symbol vector $\boldsymbol{y}=\{y_t\}$, the matrix $E_{t,s}\equiv\lambda_t^{GM}\frac{\partial \mathscr{Y}_t}{\partial r_s}$ is also quasi-Toeplitz with symbol vector $\lambda^{GM}\boldsymbol{y}=\{\lambda^{GM}y_t\}$. Therefore,

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\frac{\partial \mathscr{Y}_t}{\partial r_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}E_{t,u}=\lim_{u\to\infty}\sum_{m=-u}^{\infty}\beta^m E_{m+u,m},$$

where I have substituted $t=m+u$. Following the logic in Appendix B.1

$$\sum_{m=-u}^{\infty}\beta^m E_{m+u,m}\to\sum_{m=-\infty}^{\infty}\beta^m\lambda^{GM}y_m=\lambda^{GM}\sum_{m=-\infty}^{\infty}\beta^m y_m=\lambda^{GM}\mathscr{E}_r^{\mathscr{Y}},$$

as $u\to\infty$. Analogously, the third term goes to $\lambda^{GM}\mathscr{E}_r^{\mathscr{C}}$. This establishes that the first-order condition for the interest rate in the steady state becomes

$$\mathscr{E}_r^{\mathscr{U}}+\lambda^{GM}\left(\mathscr{E}_r^{\mathscr{Y}}-\mathscr{E}_r^{\mathscr{C}}\right)=0, \tag{23}$$

Similar reasoning can be applied to (21) and (22) to obtain the steady-state conditions for the tax rate and the progressivity of the tax system:

$$\mathscr{E}_\tau^{\mathscr{U}}+\lambda^{GM}\left(\mathscr{E}_\tau^{\mathscr{Y}}-\mathscr{E}_\tau^{\mathscr{C}}\right)=0, \tag{24}$$

$$\mathscr{E}_p^{\mathscr{U}}+\lambda^{GM}\left(\mathscr{E}_p^{\mathscr{Y}}-\mathscr{E}_p^{\mathscr{C}}\right)=0. \tag{25}$$

Using (24) to eliminate λ^{GM},

$$\mathscr{E}_r^{\mathscr{U}}=\frac{\mathscr{E}_\tau^{\mathscr{U}}}{\mathscr{E}_\tau^{\mathscr{C}}-\mathscr{E}_\tau^{\mathscr{Y}}}\left(\mathscr{E}_r^{\mathscr{C}}-\mathscr{E}_r^{\mathscr{Y}}\right), \tag{26}$$

$$\mathscr{E}_p^{\mathscr{U}}=\frac{\mathscr{E}_\tau^{\mathscr{U}}}{\mathscr{E}_\tau^{\mathscr{C}}-\mathscr{E}_\tau^{\mathscr{Y}}}\left(\mathscr{E}_p^{\mathscr{C}}-\mathscr{E}_p^{\mathscr{Y}}\right). \tag{27}$$

Equations (26) and (27), together with the goods market-clearing condition in steady state, can be used to search for a candidate Ramsey steady state $\{r, \tau, p\}$. The optimal level of debt can then be read-off the asset-market clearing condition. The advantage of this formulation is that it does not require solving for general equilibrium in each iteration. Moreover, the discounted sums of the general equilibrium responses do not enter the equations and therefore the mysterious step relying on the convolution theorem is no longer required. The disadvantage is that it requires solving for a larger system. It turns out, however, that one can use homogeneity of aggregate consumption and output to reduce the dimensionality of the system (cf. **?**). I verify that the solution implied by this formulation is consistent with the one in Appendix B.1.

B.4 Proof of Proposition 1

Recall that, for any $u = 0, 1, \ldots$, the following must be true

$$\sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial B_u}+\sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\beta^{t-u}\lambda_t\frac{\partial \mathscr{T}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial B_u}+\lambda_u-\beta\lambda_{u+1}(1+\boldsymbol{r}_u)-\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t\frac{\partial \boldsymbol{r}_t}{\partial B_u}B_{t-1}=0. \quad (28)$$

These conditions are necessary for the Ramsey plan to be optimal. If they are not satisfied, a small perturbation dB_u would increase social welfare, contradicting the optimality of the Ramsey plan. To derive (9), we let $u \to \infty$ and assume that $\lambda_t \to \lambda^{RSS} > 0$ and $B_t \to B^{RSS} < \infty$ as $t \to \infty$.

Using the quasi-Toeplitz property of Jacobians, together with the convolution theorem, as in Appendix B.1, verifies that the first term in (28) goes to $\mathscr{E}^{\mathscr{U}}_r \cdot \mathscr{E}^{\boldsymbol{r}}_B$. The third term goes to $\lambda^{RSS}\{1-\beta(1+\boldsymbol{r})\}$.

For the second term, define the matrix $M_{t,s} \equiv \lambda_t \frac{\partial \mathscr{T}_t}{\partial r_s}$. Since $\frac{\partial \mathscr{T}_t}{\partial r_s}$ is a quasi-Toeplitz matrix with some symbol vector $\boldsymbol{h} = \{h_t\}$, $M_{t,s}$ is also quasi-Toeplitz with symbol vector $\lambda^{RSS}\boldsymbol{h} = \{\lambda^{RSS}h_t\}$. Therefore,

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\beta^{t-u}\lambda_t\frac{\partial \mathscr{T}_t}{\partial r_s}\frac{\partial \boldsymbol{r}_s}{\partial B_u}=\lim_{u\to\infty}\sum_{t=0}^{\infty}\sum_{s=0}^{\infty}\beta^{t-u}M_{t,s}\frac{\partial \boldsymbol{r}_s}{\partial B_u}=\lim_{u\to\infty}\sum_{m=-u}^{\infty}\beta^m\sum_{n=-u}^{\infty}M_{m+u,n+u}\frac{\partial \boldsymbol{r}_{n+u}}{\partial B_u},$$

where I have substituted $t = m+u$ and $s = n+u$. Given the quasi-Toeplitz property of $M_{t,s}$ and $\frac{\partial \boldsymbol{r}_t}{\partial B_s}$,

$$\sum_{m=-u}^{\infty}\beta^m\sum_{n=-u}^{\infty}M_{m+u,n+u}\frac{\partial \boldsymbol{r}_{n+u}}{\partial B_u}\to\sum_{m=-\infty}^{\infty}\beta^m\sum_{n=-\infty}^{\infty}\lambda^{RSS}h_{m-n}g_n=\lambda^{RSS}\sum_{m=-\infty}^{\infty}\beta^m\sum_{n=-\infty}^{\infty}h_{m-n}g_n,$$

where $\{g_t\}$ is the symbol vector of $\frac{\partial \boldsymbol{r}_t}{\partial B_s}$. The term inside the second summation is the convolution of two summable sequences. Applying the convolution theorem and following the logic in Appendix B.1,

$$\lambda^{RSS} \sum_{m=-\infty}^{\infty} \beta^m \sum_{n=-\infty}^{\infty} h_{m-n} g_n = \lambda^{RSS} \sum_{k=-\infty}^{\infty} \beta^k h_k \sum_{k=-\infty}^{\infty} \beta^k g_k = \lambda^{RSS} \mathscr{E}_r^{\mathscr{T}} \cdot \mathscr{E}_B^{\boldsymbol{r}}$$

This establishes that the second term in (28) goes to $\lambda^{RSS} \mathscr{E}_r^{\mathscr{T}} \cdot \mathscr{E}_B^{\boldsymbol{r}}$. Analogously, fourth term in (28) goes to $\lambda^{RSS} \mathscr{E}_B^{\boldsymbol{r}} B^{RSS}$. Therefore, as $u \to \infty$, (28) becomes

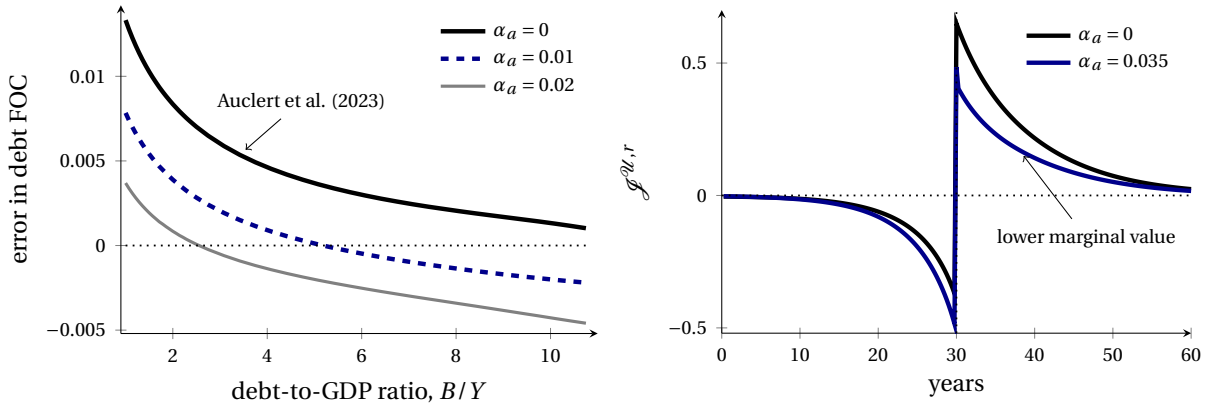
$$\mathscr{E}_r^{\mathscr{U}} \cdot \mathscr{E}_B^{\boldsymbol{r}} + \lambda^{RSS} \mathscr{E}_r^{\mathscr{T}} \cdot \mathscr{E}_B^{\boldsymbol{r}} + \lambda^{RSS} \{1 - \beta(1 + \boldsymbol{r})\} - \lambda^{RSS} \mathscr{E}_B^{\boldsymbol{r}} \, B^{RSS} = 0. \tag{29}$$

Dividing through by λ^{RSS} gives (9).

B.5 Interior solution for the Ramsey steady state

This appendix verifies that the long-run optimality conditions have an interior solution for the inequality-averse specifications considered in the paper, and explains why. Figure 12a plots the error in the first order condition for the optimal long-run choice of debt as the debt-to-GDP ratio varies, holding fixed the progressivity of the tax system. The solid line confirms the results in **?**: there is no interior root in the standard heterogeneous-agent model with separable CRRA preferences and utilitarian SWF. The dashed lines, on the other hand, indicate that the debt condition has a unique interior root for the positive values of α_a shown. The figures in Appendix B.6 provide the corresponding diagnostics when debt and progressivity are chosen jointly.

Figure 12: Interior solutions with inequality-averse generational planners



(a) The search for B^{RSS} and inequality aversion

(b) $\mathscr{E}_r^{\mathscr{U}}$ with and without inequality aversion

To understand what drives this result, I look at the Jacobians of aggregate utility to changes in the interest rate: $\mathscr{J}_{t,s}^{\mathscr{U},r} = \frac{\partial \mathscr{U}_t}{\partial r_s}$.[19] From Proposition 1, the β-discounted sum of a far-out column of $\mathscr{J}^{\mathscr{U},r}$, $\mathscr{E}_r^{\mathscr{U}} = \lim_{s\to\infty} \sum_{t=0}^{\infty} \beta^{t-s} \frac{\partial \mathscr{U}_t}{\partial r_s}$, governs the marginal social value of issuing public debt around the RSS. This is the only component in (9) that directly depends on the planner's aversion to inequality. Understanding how $\mathscr{E}_r^{\mathscr{U}}$ varies with the planner's aversion to inequality is therefore key.

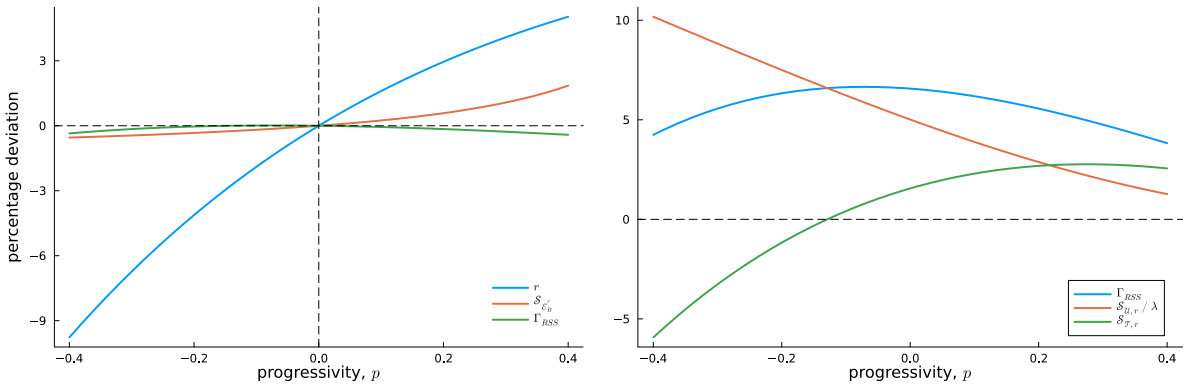
Figure 12b plots the long-run discounted response of aggregate utility for a utilitarian plan-

[19] See **?** for more on these sequence-space Jacobians.

ner (solid line) and an inequality-averse planner (dashed line). A generational planner that underweights the utility of the asset-rich assigns a lower marginal benefit to increases in public debt compared to a utilitarian planner. This difference arises for two reasons. First, the effects due to the anticipatory response, which lowers consumption and thus lowers aggregate utility prior to the shock, are stronger at the bottom of the wealth distribution: the marginal utility of consumption is higher for those who are asset-poor. Second, the benefits when the shock hits are discounted because a higher r, once realized, mostly benefits those who are asset-rich. Together, these two effects imply that the marginal benefit of increasing public debt is lower for an inequality-averse planner than for a utilitarian planner. This explains why introducing aversion to inequality restores an interior solution to the long-run Ramsey conditions.

B.6 Additional RSS figures

Figure 13: Sufficient statistics and progressivity in the RSS



(a) r, $\mathcal{E}_B^{\boldsymbol{r}}$, and Γ^{RSS} as progressivity varies

(b) Unpacking $\Gamma^{RSS} \equiv \frac{\mathcal{E}_r^{\mathcal{U}}}{\lambda^{RSS}} + \mathcal{E}_r^{\mathcal{T}}$

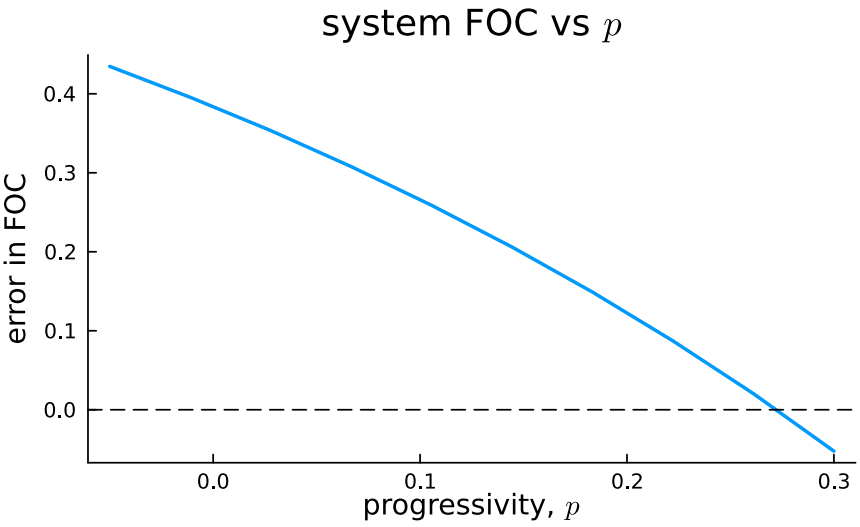


Figure 14: Verifying uniqueness of solution to (16)

C Appendix to optimal steady state problem

This appendix provides additional details on the OSS problem and its solution. I start by deriving the optimality conditions for the OSS problem. I then describe the algorithm to compute the optimal level of public debt and progressivity. I also discuss an alternative formulation of the OSS problem that allows for a sufficient-statistic representation of the optimality conditions that is similar to Proposition 1. Finally, I define the alternative social welfare functions used in the robustness exercise of Section 5.1.

C.1 Optimality conditions for OSS problem

To take into account constraints in problem (11), for each (B,p), solve for *functions* $\boldsymbol{r}(B,p)$ and $\boldsymbol{\tau}(B,p)$. Then, the OSS problem reduces to an unconstrained maximization problem:

$$\max_{B,p} \mathcal{W}(\boldsymbol{r}(B,p),\boldsymbol{\tau}(B,p),p).$$

The first-order conditions with respect to p and B are, respectively:

$$\frac{\partial \mathcal{W}}{\partial r}\cdot\frac{\partial \boldsymbol{r}}{\partial p}+\frac{\partial \mathcal{W}}{\partial p}+\frac{\partial \mathcal{W}}{\partial \tau}\cdot\frac{\partial \boldsymbol{\tau}}{\partial p}=0 \tag{30}$$

$$\frac{\partial \mathcal{W}}{\partial r}\cdot\frac{\partial \boldsymbol{r}}{\partial B}+\frac{\partial \mathcal{W}}{\partial \tau}\cdot\frac{\partial \boldsymbol{\tau}}{\partial B}=0 \tag{31}$$

To solve for the GE derivatives in (30), one can use part one of Lemma 1 in the main text. The GE derivatives in (31) can be obtained from the following system:

$$\begin{bmatrix} \frac{\partial \mathscr{A}}{\partial r} & \frac{\partial \mathscr{A}}{\partial \tau} \\ \frac{\partial \mathscr{T}}{\partial r} - B & \frac{\partial \mathscr{T}}{\partial \tau} \end{bmatrix} \begin{bmatrix} \frac{\partial \boldsymbol{r}}{\partial B} \\ \frac{\partial \boldsymbol{\tau}}{\partial B} \end{bmatrix} = \begin{bmatrix} 1 \\ r \end{bmatrix}. \tag{32}$$

I use (30) and (31) to reduce the optimization problem in (11) to a (numerical) root-finding problem.[20]

C.2 Computing the OSS

The algorithm to compute the OSS proceeds as follows:

1. Given a candidate level of public debt B and progressivity p, solve for the interest rate r and the level of taxes τ that ensure asset market clearing and government budget balance.

2. Compute the partial equilibrium derivatives of aggregate welfare $\frac{\partial \mathscr{W}}{\partial r}$, $\frac{\partial \mathscr{W}}{\partial \tau}$, $\frac{\partial \mathscr{W}}{\partial p}$, aggregate tax revenues $\frac{\partial \mathscr{T}}{\partial r}$, $\frac{\partial \mathscr{T}}{\partial \tau}$, $\frac{\partial \mathscr{T}}{\partial p}$, and aggregate asset demand $\frac{\partial \mathscr{A}}{\partial r}$, $\frac{\partial \mathscr{A}}{\partial \tau}$, $\frac{\partial \mathscr{A}}{\partial p}$ via numerical differentiation.

3. Use Lemma 1 for the general equilibrium derivatives $\frac{\partial \boldsymbol{r}}{\partial p}$ and $\frac{\partial \boldsymbol{\tau}}{\partial p}$ and (32) for $\frac{\partial \boldsymbol{r}}{\partial B}$ and $\frac{\partial \boldsymbol{\tau}}{\partial B}$.

4. Check whether (30) and (31) are satisfied. If they are, stop. Otherwise, adjust B and p and repeat the process.

As in Appendix B.2, it is better to iterate on r and p, and then use asset market clearing to read off the implied level of public debt B. This means that in step (1) above, we only need to solve for the level of taxes τ that ensures government budget balance. Once more, to avoid multi-dimensional root-finding algorithms, I separate the problem into two steps. So in step (4), I first check whether (30) is satisfied. If this is not the case, I adjust p and repeat the process but keeping r fixed. Because this is a one-dimensional problem, the updating for p can be done via Brent's method.[21] Once (30) is satisfied at the candidate level of r, I check whether (31) is

[20] I check that the solution is unique, so there is no need to worry about multiple roots. Even if this were the case, one could easily rank them by evaluating the objective function because doing so is not computationally hard.

[21] Here, I also check that there is a unique value of p that solves (30).

satisfied. If this is not the case, I adjust r and repeat the process, resolving for p along the way. Finally, note that step (3) does not require re-calculating the equilibrium, as the GE derivatives can be expressed in terms of PE derivatives.

C.3 Alternative OSS formulation

To avoid solving for general equilibrium in each iteration, one can also work with the goods market clearing condition. Indeed, by Walras' Law, the OSS problem in (11) is equivalent to:

$$\max_{\{r,\tau,p,B\}} \mathscr{W}(r,\tau,p) \quad \text{s.t} \quad \begin{cases} \mathscr{A}(r,\tau,p) = B, \\ \mathscr{C}(r,\tau,p) + G = \mathscr{Y}(r,\tau,p) \end{cases} . \tag{33}$$

Notice that, given an interest rate r and a CRP tax code $\{\tau, p\}$, the planner can always find a level of public debt B to ensure asset market clearing holds. After dropping this constraint and the associated choice variable, the problem reduces to

$$\max_{\{r,\tau,p\}} \mathscr{W}(r,\tau,p) \quad \text{s.t} \quad \mathscr{C}(r,\tau,p) + G = \mathscr{Y}(r,\tau,p).$$

The Lagrangian for this problem is

$$\max_{\{r,\tau,p\}} \mathscr{W}(r,\tau,p) + \lambda^{GM}\left\{\mathscr{Y}(r,\tau,p) - \mathscr{C}(r,\tau,p) - G\right\}.$$

The associated first-order conditions are

$$\begin{aligned} \frac{\partial \mathscr{W}}{\partial \tau} + \lambda^{GM}\left\{\frac{\partial \mathscr{Y}}{\partial \tau} - \frac{\partial \mathscr{C}}{\partial \tau}\right\} = 0, \\ \frac{\partial \mathscr{W}}{\partial r} + \lambda^{GM}\left\{\frac{\partial \mathscr{Y}}{\partial r} - \frac{\partial \mathscr{C}}{\partial r}\right\} = 0, \\ \frac{\partial \mathscr{W}}{\partial p} + \lambda^{GM}\left\{\frac{\partial \mathscr{Y}}{\partial p} - \frac{\partial \mathscr{C}}{\partial p}\right\} = 0, \end{aligned}$$

together with the goods market clearing condition. Using the first one to eliminate λ^{GM}, I arrive at

$$\frac{\partial \mathscr{W}}{\partial r} = \frac{\frac{\partial \mathscr{W}}{\partial \tau}}{\frac{\partial \mathscr{C}}{\partial \tau} - \frac{\partial \mathscr{Y}}{\partial \tau}}\left\{\frac{\partial \mathscr{C}}{\partial r} - \frac{\partial \mathscr{Y}}{\partial r}\right\} \tag{34}$$

$$\frac{\partial \mathscr{W}}{\partial p} = \frac{\frac{\partial \mathscr{W}}{\partial \tau}}{\frac{\partial \mathscr{C}}{\partial \tau} - \frac{\partial \mathscr{Y}}{\partial \tau}}\left\{\frac{\partial \mathscr{C}}{\partial p} - \frac{\partial \mathscr{Y}}{\partial p}\right\} \tag{35}$$

Equations (34) and (35), combined with the goods market clearing condition, can be used to solve for the three unknowns $\{r, \tau, p\}$. As before, the advantage of this formulation is that it does not require solving for general equilibrium in each iteration. I use homogeneity of aggregate consumption and output to reduce the dimensionality of the system and verify that the solution implied by this formulation is consistent with the solution obtained via the algorithm in C.2.

C.4 Sufficient statistic representation of optimality conditions

I derive a simple "sufficient-statistic" representation for the optimal choice of debt B that is analogous to the one derived in Proposition 1. Fix p and τ, and zoom-in on the optimal choice of B. Let $\boldsymbol{r}(B, \tau, p)$ denote the interest rate that clears the asset market given the fiscal policy of the government. Then, letting λ denote the Lagrange multiplier on the government's budget constraint, the optimal choice of debt is

$$B^{OSS} = \arg\max_{B} \left\{ \mathscr{W}(\boldsymbol{r}(B, \tau, p), \tau, p) + \lambda \left(\mathscr{T}(\boldsymbol{r}(B, \tau, p), \tau, p) - \boldsymbol{r}(B, \tau, p) B - G \right) \right\}. \tag{36}$$

At an interior solution, the first order condition for this problem implies

$$\mathscr{W}_r \frac{\partial \boldsymbol{r}}{\partial B} + \lambda \left\{ \mathscr{T}_r \frac{\partial \boldsymbol{r}}{\partial B} - \frac{\partial \boldsymbol{r}}{\partial B} B^{OSS} - \boldsymbol{r} \right\} = 0. \tag{37}$$

Define the social marginal value of public debt, in dollar terms, as $\Gamma \equiv \frac{\mathscr{W}_r}{\lambda} + \mathscr{T}_r$. Similarly, let $\epsilon_B^r \equiv \frac{\partial \boldsymbol{r}}{\partial B} \frac{B^{OSS}}{1+\boldsymbol{r}}$ denote the elasticity of interest rates with respect to changes in the level of public debt *at the optimum*. Then, (37) can be written as

$$\Gamma \cdot \epsilon_B^r - \frac{\boldsymbol{r}}{1+\boldsymbol{r}} B^{OSS} - \epsilon_B^r \cdot B^{OSS} = 0.$$

Solving for B^{OSS} yields:

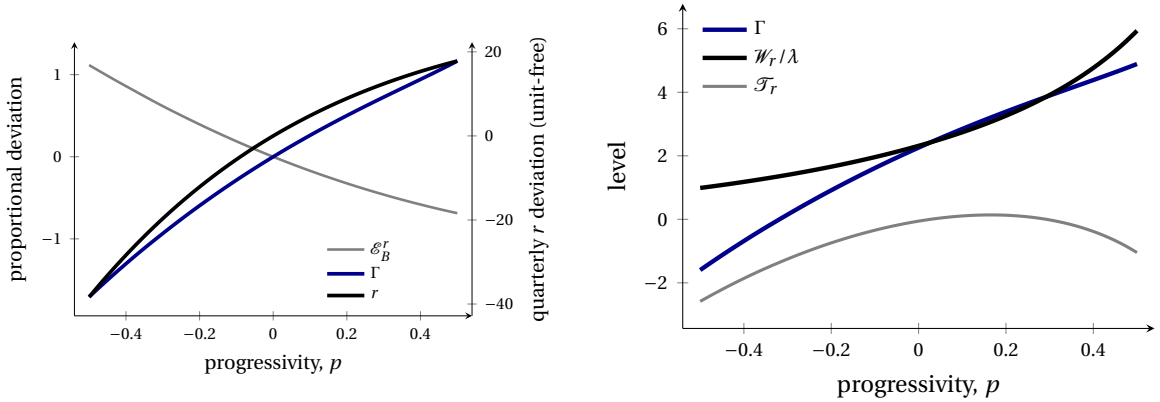
$$B^{OSS} = \frac{\epsilon_B^r}{\epsilon_B^r + \frac{\boldsymbol{r}}{1+\boldsymbol{r}}} \times \Gamma. \tag{38}$$

This equation shows that the optimal level of public debt depends on three objects: the social marginal value of public debt Γ, the *elasticity* of interest rates with respect to changes in the level of public debt $\mathscr{E}_B^r$, and the *level of interest rates* $\boldsymbol{r}$. This is similar to the sufficient-statistic representation derived in Proposition 1. The difference, however, is that the premium no longer appears in this equation. Intuitively, this is because debt only enters as a cost in the optimal steady state problem.

A higher interest rate unambiguously decreases the optimal level of public debt, holding everything else equal. Somewhat more subtle, a higher elasticity $\mathscr{E}_B^r$ increases B^{OSS} when interest rates are positive. This is not obvious, since this object affects both the costs and benefits of public debt. But at an interior optimum, if $\boldsymbol{r} \geq 0$, it must be that $\Gamma \geq B^{OSS}$. Therefore, a more elastic interest rate increases the marginal benefit of public debt by more than it increases the fiscal cost, pushing towards higher B^{OSS}. Finally, when the social marginal value of public debt Γ increases, the optimal level of public debt also increases, holding everything else fixed.

This discussion implies that the progressivity of the tax system affects the optimal level of public debt through three channels: $\boldsymbol{r}$, $\mathscr{E}_B^r$, and Γ. The interest-rate channel, emphasized throughout the paper, implies that a more progressive tax system increases $\boldsymbol{r}$ and hence lowers the optimal level of public debt. Pushing in the same direction, a more progressive p lowers the marginal *private value* of B, as it helps agents insure against risk and thus reduces the value of public liquidity. This reinforces the effects driven by the interest rate channel. With income effects, however, the response of the marginal *social value* Γ, that includes revenue effects, can be less obvious. The effect of p on $\mathscr{E}_B^r$ is more complex. But this elasticity is very low in the model (compared to its data counterpart) and not too responsive to changes in p. So this channel does not seem to play a major role.

Figure 15: Sufficient statistics and progressivity in the OSS



(a) r, $\mathscr{E}_B^r$, and Γ as progressivity varies

(b) Unpacking Γ as progressivity varies

C.5 Alternative social welfare functions

This appendix states the two alternative SWFs used in Section 5.1. The first is a generalized utilitarian criterion with weights that depend on asset holdings. Formally,

$$\mathcal{W}^{GU}(r,\tau,p) = \int \omega(a_i) V(\theta_i, a_i)\, di, \tag{39}$$

with the weights parameterized as in the main text. Relative to a generational planner, this kind of social planner only cares about inequality of the initial generation: the weight on the utility of future generations does not depend on the asset holdings of that generation. In this sense, the social preferences behind equation (7) may reflect a stronger notion of aversion to inequality than the ones behind equation (39).

The second, following **?** and **?**, separates society's aversion to inequality from household's aversion to intertemporal fluctuations. Letting $\bar{c}_i$ denote the consumption certainty-equivalent of agent i[22], social welfare for a "Bénabou planner" is given by

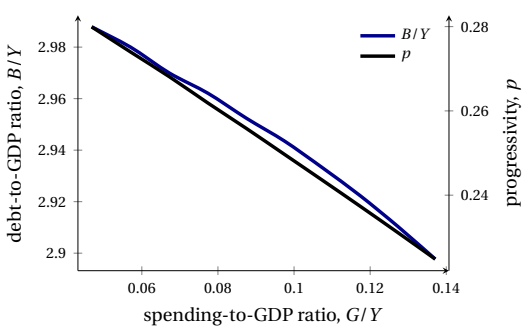
$$\mathcal{W}^{\alpha}(r,\tau,p) = \left(\int \bar{c}_i^{1-\frac{1}{\alpha}}\, di \right)^{\frac{1}{1-\frac{1}{\alpha}}}. \tag{40}$$

The parameter $\alpha \in (0,\infty)$ governs society's aversion to inequality. When $\alpha \to \infty$, the objective captures pure economic efficiency and puts no value on equity of consumption per se – redistribution has value only to the extent that it relaxes borrowing constraints or reduces idiosyncratic risk. As **?** puts it, efficiency concerns are thus separated from pure equity concerns. If α equals the agents' elasticity of intertemporal substitution, one recovers the standard utilitarian criterion. As $\alpha \to 0$, the objective reduces to that of a Rawlsian planner who only cares about the welfare of the poorest agents.

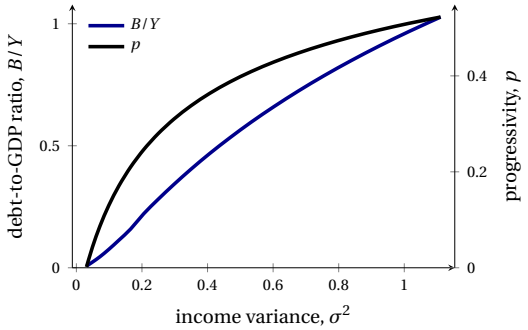
[22] Formally, $\bar{c}_i$ is the constant level of consumption a household would need to receive, without working, in order to achieve the equilibrium lifetime utility $V(\theta_i, a_i)$.

C.6 Additional OSS figures

Figure 16: Comparative statics with respect to spending and income risk



(a) Optimal mix and government spending



(b) Optimal mix and income risk

D Appendix to model with capital

This appendix provides details on the model with capital. Appendix D.1 defines the equilibrium for this economy and describes the calibration of the model with capital. Appendix D.2 presents the dynamic Ramsey problem in the model with capital and proves the optimality of the modified golden rule. Finally, Appendix D.4 derives the full set of optimality conditions and details the computational procedure to solve for the Ramsey steady state in the model with capital.

D.1 Equilibrium definition and calibration for model with capital

Definition 2 An equilibrium in the model with capital is a sequence of prices $\{\bar{r}_t, \bar{w}_t, r_t, w_t\}$ and tax schedules $\{p_t, \tau_t, \tau_t^k\}$, a sequence of policy functions $\{\boldsymbol{c}_t(x), \boldsymbol{a}_t(x), \boldsymbol{y}_t(x)\}$, aggregates $\{L_t, K_t, B_t\}$, and distributions $\{D_t\}$ such that, given the initial distribution D_0 and the initial level of capital K_0, the following conditions hold:

(i) $\{\boldsymbol{c}_t(x), \boldsymbol{a}_t(x), \boldsymbol{y}_t(x)\}$ are optimal given $\{\bar{r}_t, \bar{w}_t\}$,

(ii) D_t is consistent with the policy functions and the Markov process for productivity,

(iii) after-tax wages $\bar{w}_t = \tau_t w_t^{1-p_t}$ and after-tax interest rates $\bar{r}_t = (1-\tau_{kt}) r_t$,

(iv) before-tax wages $w_t = F_L(K_{t-1}, L_t)$ and before-tax interest rates $r_t = F_K(K_{t-1}, L_t) - \delta$,

(v) the government's budget constraint is satisfied

$$G + (1+r_t)B_{t-1} = B_t + \mathscr{T}_t(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}) + \tau_t^k \mathscr{A}_{t-1}(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}),$$

(vi) the asset market, the labor market, and the goods market clear

$$\mathscr{A}_t\left(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}\right) = B_t + K_t,$$
$$\mathscr{L}_t(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}) = L_t,$$
$$\mathscr{C}_t\left(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}\right) + K_t + G = F(K_{t-1}, L_t) + (1-\delta)K_{t-1}.$$ □

The model with multiple safe assets is once again calibrated to the US economy. The capital share α is chosen in order to generate a capital-to-GDP ratio of 2.5, in line with US data. The

discount factor β is then chosen to match a real interest rate of 2%, given the capital-to-GDP ratio of 2.5 and a debt-to-GDP ratio of 0.615. This is based on the average US federal debt in the data from 1995 to 2007.[23] The parameters of the income process and the borrowing limit are unchanged relative to Section 2. The depreciation rate for capital is 2% at the quarterly frequency. The value for the tax on capital income is taken from **?**, whose estimation for the US in 2007 yields 36%. Government spending and the progressivity of the tax system are unchanged. Finally, the level of labor income taxes τ adjusts in order to ensure that the government budget constraint holds. Table 2 summarizes the parameter values that come out from this procedure.

Table 2: Parameter values in model with multiple safe assets

Parameter	*Description*	*Value*	*Parameter*	*Description*	*Value*
β	discounting	0.993	G/Y	spending-to-GDP	0.088
ρ	persistence of AR (1)	0.966	K/Y	capital-to-GDP	2.5
σ_ϵ^2	shock variance	0.033	B/Y	debt-to-GDP	0.615
EIS	curvature in u	1	p	tax progressivity	0.181
$Frisch$	curvature in v	1/2	τ	level of taxes	1.017
α	capital share	0.25	τ_k	capital income tax	0.36
δ	depreciation rate	0.02	ϕ	borrowing limit	0

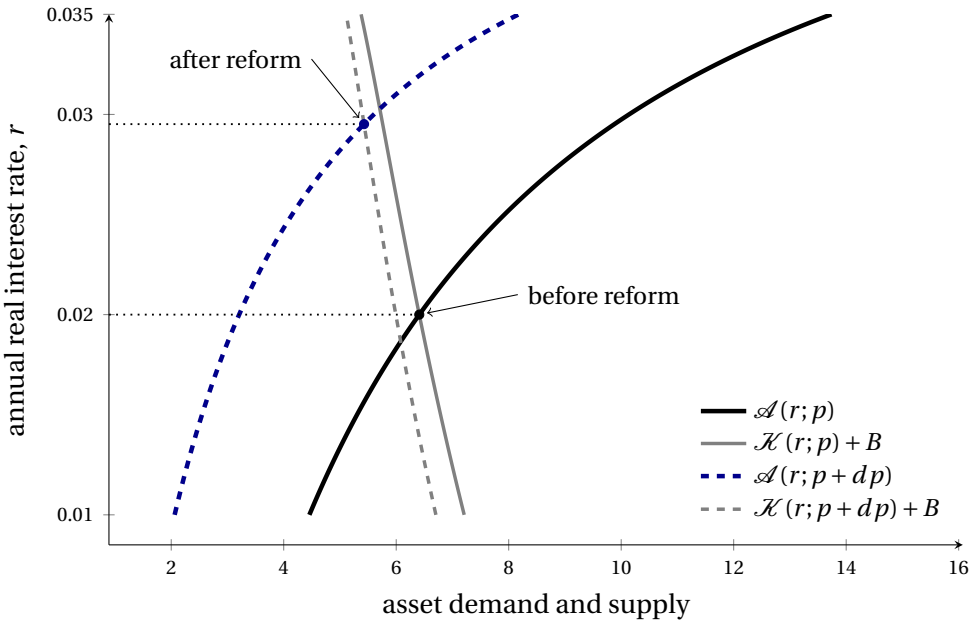
Figure 17 repeats the asset market diagram of Figure 1 for this economy. The supply of safe assets now slopes downward, since a higher interest rate lowers the desired capital stock. A progressive tax reform therefore moves the equilibrium interest rate by less than in the baseline model.

D.2 Ramsey problem in the model with capital and τ_k

In the model with capital, the dynamic Ramsey problem can be written as

[23] This is based on **?** and is close to the values in **?** and **?**.

Figure 17: The interest rate channel in the model with capital



Note: The black curves are asset demand before (solid) and after (dashed) a progressive tax reform. The grey curves are the supply of safe assets, $\mathscr{K}(r;p)+B$, before and after the reform. The dotted lines mark the equilibrium interest rates.

$$\max_{\{\bar{r}_t,\bar{w}_t,r_t,p_t,B_t,K_t,L_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{\bar{r}_s\},\{\bar{w}_s\},\{p_s\}\right) \quad \text{s.t} \begin{cases} \mathscr{A}_t\left(\{\bar{r}_s\},\{\bar{w}_s\},\{p_s\}\right) = B_t + K_t, \\ \mathscr{C}_t\left(\{\bar{r}_s\},\{\bar{w}_s\},\{p_s\}\right) + K_t + G = F(K_{t-1},L_t) + (1-\delta)K_{t-1}, \\ w_t = F_L(K_{t-1},L_t),\ r_t = F_K(K_{t-1},L_t) - \delta, \\ \mathscr{L}_t(\{\bar{r}_s\},\{\bar{w}_s\},\{p_s\}) = L_t \end{cases}$$

Once again, $\bar{r}_t = (1-\tau_{kt})r_t$ and $\bar{w}_t = \tau_t w_t^{1-p_t}$ denote the after-tax interest rate and the after-tax wage. Notice that we drop the government budget constraint because of Walras' Law. As in the main text, $\mathscr{U}_t$, $\mathscr{A}_t$, $\mathscr{C}_t$, and $\mathscr{L}_t$ are *sequence-space* functions that map sequences of after-tax interest rates, after-tax wages and progressivity into aggregates at time t. Since these functions aggregate optimal individual behavior using the distribution of agents, the Ramsey plan is implementable, provided that the constraints listed above are satisfied.

D.3 Proving the optimality of the modified golden rule

Given choices of $\{\bar{r}_t, \bar{w}_t, p_t, K_t\}$, asset market clearing condition pins down $\{B_t\}$. Similarly, given these choices, the labor market clearing condition pins down the sequence of labor demand $\{L_t\}$. The firm's optimality conditions then pin down the sequence of pre-tax wages $\{w_t\}$ and pre-tax interest rates $\{r_t\}$. Thus, after dropping the redundant choice variables and associated constraints, the RSS problem reduces to

$$\max_{\{\bar{r}_t, \bar{w}_t, p_t, K_t,\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \mathscr{U}_t\left(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}\right) \quad \text{s.t} \quad \mathscr{C}_t + K_t + G = F(K_{t-1}, \mathscr{L}_t) + (1-\delta)K_{t-1}$$

The Lagrangian for this problem is

$$\max_{\{\bar{r}_t, \bar{w}_t, p_t, K_t,\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left\{\mathscr{U}_t\left(\{\bar{r}_s\}, \{\bar{w}_s\}, \{p_s\}\right) + \lambda_t^{GM}\left\{F(K_{t-1}, \mathscr{L}_t) + (1-\delta)K_{t-1} - \mathscr{C}_t - K_t - G\right\}\right\}$$

The first-order condition with respect to capital is

$$\lambda_t^{GM} = \beta \lambda_{t+1}^{GM}\left[F_K\left(K_t, \mathscr{L}_{t+1}(\cdot)\right) + (1-\delta)\right], \quad \text{for } t = 0, 1, \ldots. \tag{41}$$

Using homogeneity of degree one of the production function and defining the capital labor ratio $k_t \equiv \frac{K_t}{\mathscr{L}_{t+1}(\cdot)}$, we can rewrite this first order condition as

$$\lambda_t^{GM} = \beta \lambda_{t+1}^{GM}\left[F_K\left(k_t, 1\right) + (1-\delta)\right], \quad \forall\, t.$$

If quantities and multipliers converge to an interior steady state, this condition becomes

$$1 = \beta\left[F_K(k, 1) + (1-\delta)\right]$$

This establishes the optimality of the modified golden rule. This result was first proved by **?** under the assumption that government spending is endogenous. Here, I verify that it goes through in the model with exogenous government spending and regardless of the planner's preference for redistribution. As discussed by **?**, two situations would prevent the applicability of this result: (i) nonconvergence to an interior steady state and (ii) nonconvergence of the multipliers. These are ruled out by assumption.

D.4 Additional first-order conditions

In addition to the first-order condition with respect to capital (41), we also have:

$$\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial \bar{r}_u}+\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\left(w_t\frac{\partial \mathscr{L}_t}{\partial \bar{r}_u}-\frac{\partial \mathscr{C}_t}{\partial \bar{r}_u}\right)=0,\quad \text{for } u=0,1,\ldots \tag{42}$$

$$\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial \bar{w}_u}+\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\left(w_t\frac{\partial \mathscr{L}_t}{\partial \bar{w}_u}-\frac{\partial \mathscr{C}_t}{\partial \bar{w}_u}\right)=0,\quad \text{for } u=0,1,\ldots \tag{43}$$

$$\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{U}_t}{\partial p_u}+\sum_{t=0}^{\infty}\beta^{t-u}\lambda_t^{GM}\left(w_t\frac{\partial \mathscr{L}_t}{\partial p_u}-\frac{\partial \mathscr{C}_t}{\partial p_u}\right)=0,\quad \text{for } u=0,1,\ldots \tag{44}$$

If quantities and multipliers converge to an interior steady state, the same logic as in Appendix B.3 verifies that when $u\to\infty$, these conditions become

$$\mathscr{E}_{\bar{r}}^{\mathscr{U}}+\lambda^{GM}\left(w^{MGR}\mathscr{E}_{\bar{r}}^{\mathscr{L}}-\mathscr{E}_{\bar{r}}^{\mathscr{C}}\right)=0, \tag{45}$$

$$\mathscr{E}_{\bar{w}}^{\mathscr{U}}+\lambda^{GM}\left(w^{MGR}\mathscr{E}_{\bar{w}}^{\mathscr{L}}-\mathscr{E}_{\bar{w}}^{\mathscr{C}}\right)=0, \tag{46}$$

$$\mathscr{E}_{p}^{\mathscr{U}}+\lambda^{GM}\left(w^{MGR}\mathscr{E}_{p}^{\mathscr{L}}-\mathscr{E}_{p}^{\mathscr{C}}\right)=0, \tag{47}$$

where $w^{MGR}=F_L(k^{MGR},1)$ is the pre-tax wage implied by the modified golden rule. One can use the three conditions above, after imposing the MGR, together with the goods market clearing condition to solve for the unknowns $\{\bar{r},\bar{w},p,\lambda^{GM}\}$. Appendix D.5 details the computational procedure.

D.5 Computing the RSS in the model with capital

To operationalize (45)-(47) one needs to compute the discounted sum of the asymptotic response of aggregate outcomes $\mathscr{Y}\in\{\mathscr{U},\mathscr{L},\mathscr{C}\}$ to changes in the instruments of the planner:

$$\lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{Y}_t}{\partial \bar{r}_u},\quad \lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{Y}_t}{\partial \bar{w}_u},\quad \lim_{u\to\infty}\sum_{t=0}^{\infty}\beta^{t-u}\frac{\partial \mathscr{Y}_t}{\partial p_u}.$$

This could be done by computing the Jacobians for welfare, labour supply, and consumption, around the steady state implied by some candidate $(\bar{r},\bar{w},p)$ and then taking the discounted sum of some far out column. Even if one uses the methods in **?**, this turns out to be somewhat costly since we need a pretty large horizon to get convergence of the discounted sum. Therefore, I follow the same approach as in Appendix B.2 to compute the discounted long-run responses.

With these in hand, I solve for the unknowns $\{\bar{r}, \bar{w}, p, \lambda^{GM}\}$ as follows. First, I reduce the dimensionality of the system by using (46) to substitute out λ^{GM}. This leaves (45) and (47), together with goods-market clearing, as a system for $\{\bar{r}, \bar{w}, p\}$. Using homogeneity of aggregate consumption and aggregate labor supply, I can solve for $\bar{w}$ from the goods market clearing condition, reducing the system to two equations. I then follow the same two-step procedure as in Appendix B.2 to solve for $\{\bar{r}, p\}$.

E Appendix: Computing the optimal transition

This appendix describes the algorithm to compute the optimal transition path. The approach combines the sequence-space methods used to characterize the Ramsey steady state with policy gradient techniques adapted from machine learning.

E.1 Overview

The Ramsey problem is stated in Section 4. As shown in Appendix B.3, by Walras' Law this problem can be reformulated with goods market clearing as the only constraint, after dropping the government budget constraint and the associated choice variable B_t. The transition algorithm works with this equivalent formulation. The key computational challenge is the high dimensionality of the problem, as the planner must optimize over policy sequences of length T.

To solve this, I combine insights from two existing approaches. **?** introduce sequence-space Jacobians to characterize the Ramsey steady state, but do not provide a general method for global optimization over transition paths. **?** optimize over parameterized paths in the time domain, but use computationally expensive brute-force methods.

I integrate these methods by parameterizing policy paths in the time domain, following **?**, and using the sequence-space structure to compute analytical gradients of an augmented Lagrangian. The key extension relative to **?** is computing the Jacobians around arbitrary transition paths rather than a steady state.

E.2 Augmented Lagrangian formulation

Let $x = \{r_t, p_t, \tau_t\}_{t=0}^T$ denote the vector of all policy variables stacked over time, and let $\boldsymbol{g}(x) = \{\mathscr{Y}_t - \mathscr{C}_t - G\}_{t=0}^T$ denote the goods-market residuals. The augmented Lagrangian is:

$$\mathscr{L}_{\rho_{AL}}(x, \lambda) = -W(x) + \lambda^\top \boldsymbol{g}(x) + \frac{\rho_{AL}}{2} \|\boldsymbol{g}(x)\|^2, \tag{48}$$

where $W(x) = \sum_{t=0}^T \beta^t \mathscr{U}_t$ is aggregate welfare, $\lambda \in \mathbb{R}^T$ is the vector of Lagrange multipliers, and $\rho_{AL} > 0$ is a penalty parameter.

The gradient of the augmented Lagrangian with respect to the policy path is:

$$\nabla_x \mathscr{L}_{\rho_{AL}} = -\nabla_x W + (\nabla_x \boldsymbol{g})^\top (\lambda + \rho_{AL} \boldsymbol{g}), \tag{49}$$

which can be expressed in terms of sequence-space Jacobians as:

$$\nabla_x \mathscr{L}_{\rho_{AL}} = -\boldsymbol{J}_U^\top \boldsymbol{\beta} + (\boldsymbol{J}_Y - \boldsymbol{J}_C)^\top (\lambda + \rho_{AL} \boldsymbol{g}), \tag{50}$$

where $\boldsymbol{J}_U$, $\boldsymbol{J}_Y$, and $\boldsymbol{J}_C$ are the $T \times 3T$ Jacobians of aggregate utility, output, and consumption with respect to the policy path x (i.e., the entries of $\boldsymbol{J}_U$ are $\frac{\partial \mathscr{U}_t}{\partial x_s}$), and $\boldsymbol{\beta} = (1, \beta, \ldots, \beta^T)^\top$ is the vector of discount factors.

E.3 Policy parameterization

To reduce dimensionality, I parameterize policy paths as:

$$x_t^j(\theta) = \bar{x}^j + e_t \sum_{k=1}^{K_j} \theta_k^j \phi_k(t), \tag{51}$$

where $\bar{x}^j$ is the terminal Ramsey steady-state value for instrument $j \in \{r, p, \tau\}$, $\{\phi_k\}_{k=1}^K$ are Chebyshev basis functions, and $e_t = (1 - t/T)^\kappa$ is an envelope that forces convergence to the terminal state. In practice, I use $K_j \approx 3$–5 basis functions per instrument. The baseline results in Section 4 use $\kappa = 10$. Section E.6 below examines sensitivity to this choice.

The gradient with respect to parameters follows from the chain rule:

$$\nabla_\theta \mathscr{L}_{\rho_{AL}} = \left(\frac{\partial x}{\partial \theta}\right)^\top \nabla_x \mathscr{L}_{\rho_{AL}} = \Phi^\top (e \odot \nabla_x \mathscr{L}_{\rho_{AL}}), \tag{52}$$

where Φ is the $T \times K$ basis matrix and e is the envelope vector.

To ensure policy paths remain within sensible bounds (e.g., $r_t > r_{\min}$), I apply a smooth transformation:

$$x_t^j = x_{\min}^j + \frac{x_{\max}^j - x_{\min}^j}{2}\left(1 + \tanh(z_t^j)\right), \tag{53}$$

where z_t^j is the unconstrained path. The gradient is adjusted accordingly using the derivative of the tanh transformation.

E.4 Efficient Jacobian computation

While **?** show how to compute Jacobians efficiently around a steady state using the "fake news" algorithm, expanding this to arbitrary transition paths requires care. The standard algorithm relies on the fact that, at the steady state, expectations of future marginal values are time-invariant. Along a transition path, these objects change over time.

I generalize the approach by noting that the recursive structure of the household problem still applies: a shock at time s does not affect optimal decisions at time $t > s$. Thus, to compute the s-th column of the Jacobian, the backward iteration need only start from the unperturbed continuation value at $t = s+1$.

Parallelization. A key ingredient for feasibility is the parallelization of the Jacobian construction. Since the effect of a shock at time s on outcomes is independent of shocks at other times (for the purpose of constructing the Jacobian column), I compute each column $s \in \{0, \ldots, T\}$ in parallel. This leads to near-linear scaling with the number of cores and reduces the total computation time by a factor of 5 on standard workstations.

E.5 Optimization procedure

The outer loop uses the standard augmented Lagrangian method:

1. Initialize $\theta = 0$, $\lambda = 0$, $\rho_{AL} = 1$
2. **Inner loop**: Minimize $\mathscr{L}_{\rho_{AL}}(\theta, \lambda)$ over θ using L-BFGS with the analytical gradient (50). Continue until stationarity tolerance is met or a maximum number of iterations is reached.
3. **Multiplier update**: $\lambda \leftarrow \lambda + \rho_{AL} \cdot g(x^*)$

4. **Penalty update**: If constraint violation $\|g\|_\infty$ has not decreased sufficiently, increase the penalty: $\rho_{AL} \leftarrow 5\rho_{AL}$

5. Repeat steps 2–4 until $\|g\|_\infty < \varepsilon_g$ and $\|\nabla_\theta \mathcal{L}_{\rho_{AL}}\|_\infty < \varepsilon_\theta$

E.6 Sensitivity analysis

A key hyperparameter is κ, the curvature of the envelope function $e_t = (1 - t/T)^\kappa$ that anchors the policy path to the steady state. The baseline results in Section 4 use $\kappa = 10$.

Figure 18 plots the optimal transition paths under $\kappa \in \{5, 6, 8\}$. The qualitative patterns are the same across specifications: the planner flattens the tax schedule on impact, interest rates fall, and debt gradually rises towards the long-run target. There are quantitative differences but the overall shape of the transition is stable.

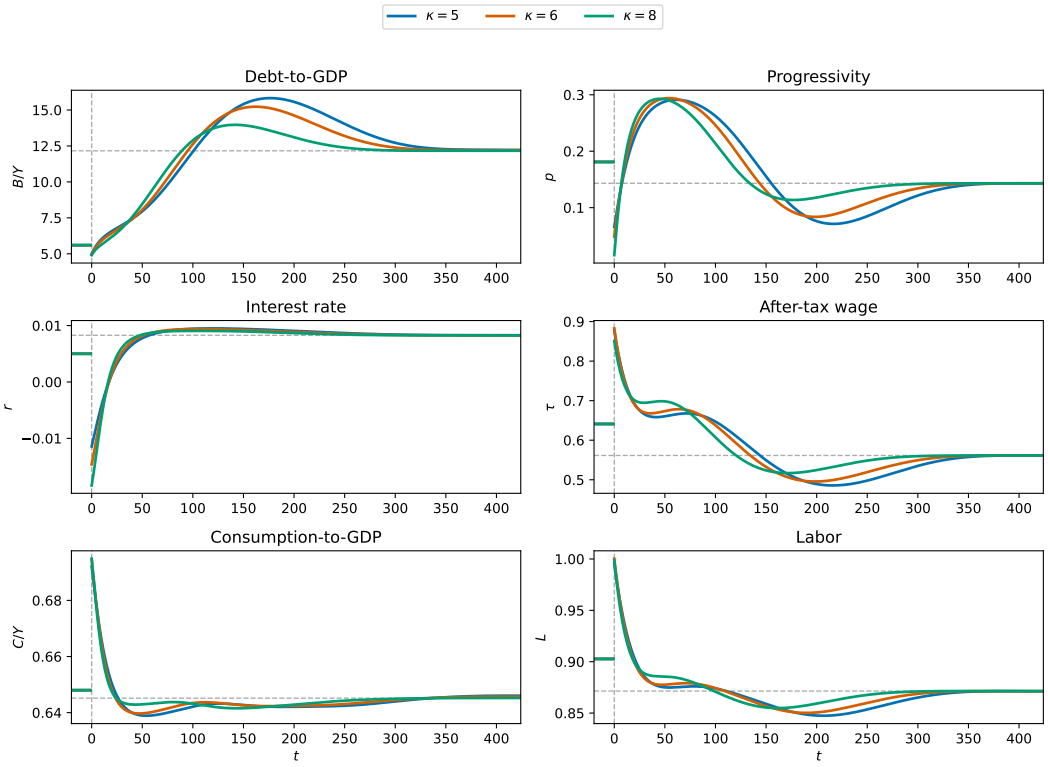
Figure 19 reports welfare gains across κ, measured as a permanent consumption-equivalent variation (CEV) relative to the status quo. Welfare is increasing in κ but the gains flatten out for $\kappa \geq 8$, suggesting that the baseline choice of $\kappa = 10$ is in the region where the approximation is not binding.

Sensitivity to basis degree. Figure 20 compares the optimal transition paths under degree-4 and degree-5 Chebyshev parameterizations, holding $\kappa = 10$ fixed. The paths are close, suggesting that the baseline results are not sensitive to increases in the dimensionality of the approximation.

E.7 Beyond smooth paths

The parameterization restricts the planner to smooth deviations around the Ramsey steady state. This restriction is not innocuous. In exploratory experiments with richer path parameterizations, the optimum moves toward front-loaded confiscation: interest rates pinned at their lower bound for the first two years and tax adjustments that saturate any imposed smoothness caps in the first decade—the classic capital-levy logic, operating here through surprise debt revaluation. Welfare gains from such paths scale with the imposed bounds rather than with model fundamentals: a feasible bound-saturating path attains a CEV of +2.2% against +1.5%

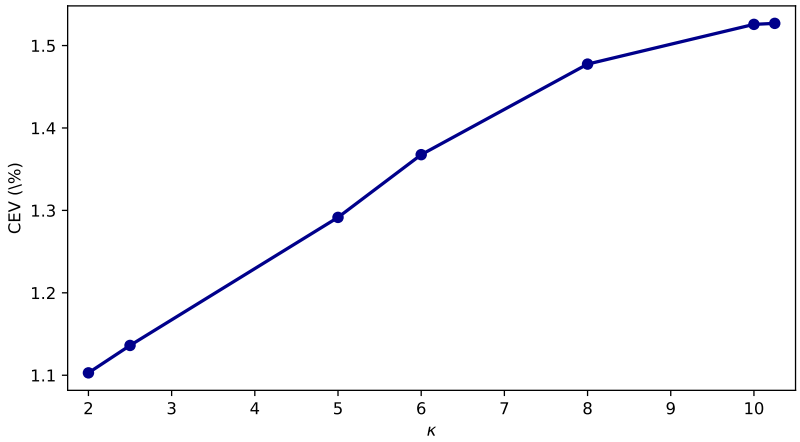
Figure 18: Sensitivity of optimal transition to envelope parameter κ



Note: Policy paths for $\{r_t, p_t, \tau_t\}$ are parameterized over $T_{\text{free}} = 425$ quarters using a Chebyshev time basis (degree 4 per instrument) and an envelope $e_t = (1 - t/T_{\text{free}})^{\kappa}$. Each line corresponds to a converged run that varies only κ (all other settings held fixed). The transition appends a constant tail of length $T_{\text{tail}} = 375$ at terminal policies. Pre-reform periods at the initial steady state are shown to the left of $t = 0$; dashed lines indicate Ramsey steady-state values.

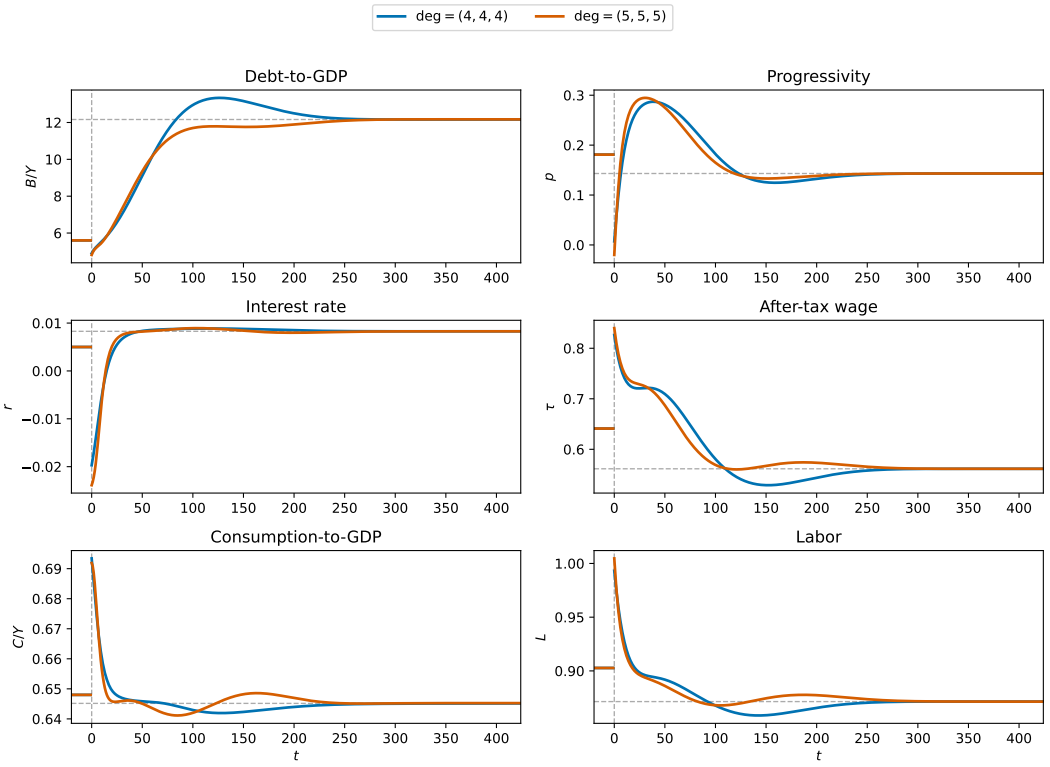
for the smooth benchmark under the baseline interest-rate floor. I therefore restrict attention to smooth implementable paths, and leave the front-loading margin to the model with capital, where a capital tax bounds it economically.

Figure 19: Welfare sensitivity to envelope parameter κ (CEV)



Note: Welfare is reported as a permanent consumption-equivalent variation (CEV) relative to the status quo policy that holds instruments fixed at their initial steady-state values. For each κ, welfare is computed from the converged transition with $T_{\text{free}} = 425$ quarters and a constant tail of length $T_{\text{tail}} = 375$ at terminal policies.

Figure 20: Sensitivity of optimal transition to Chebyshev degree



Note: Both runs use $\kappa = 10$ and $T_{\text{free}} = 425$ with a tail of $T_{\text{tail}} = 375$. The degree-5 run is warm-started from the degree-4 solution. Dashed lines indicate Ramsey steady-state values.

F Appendix to cross-country evidence

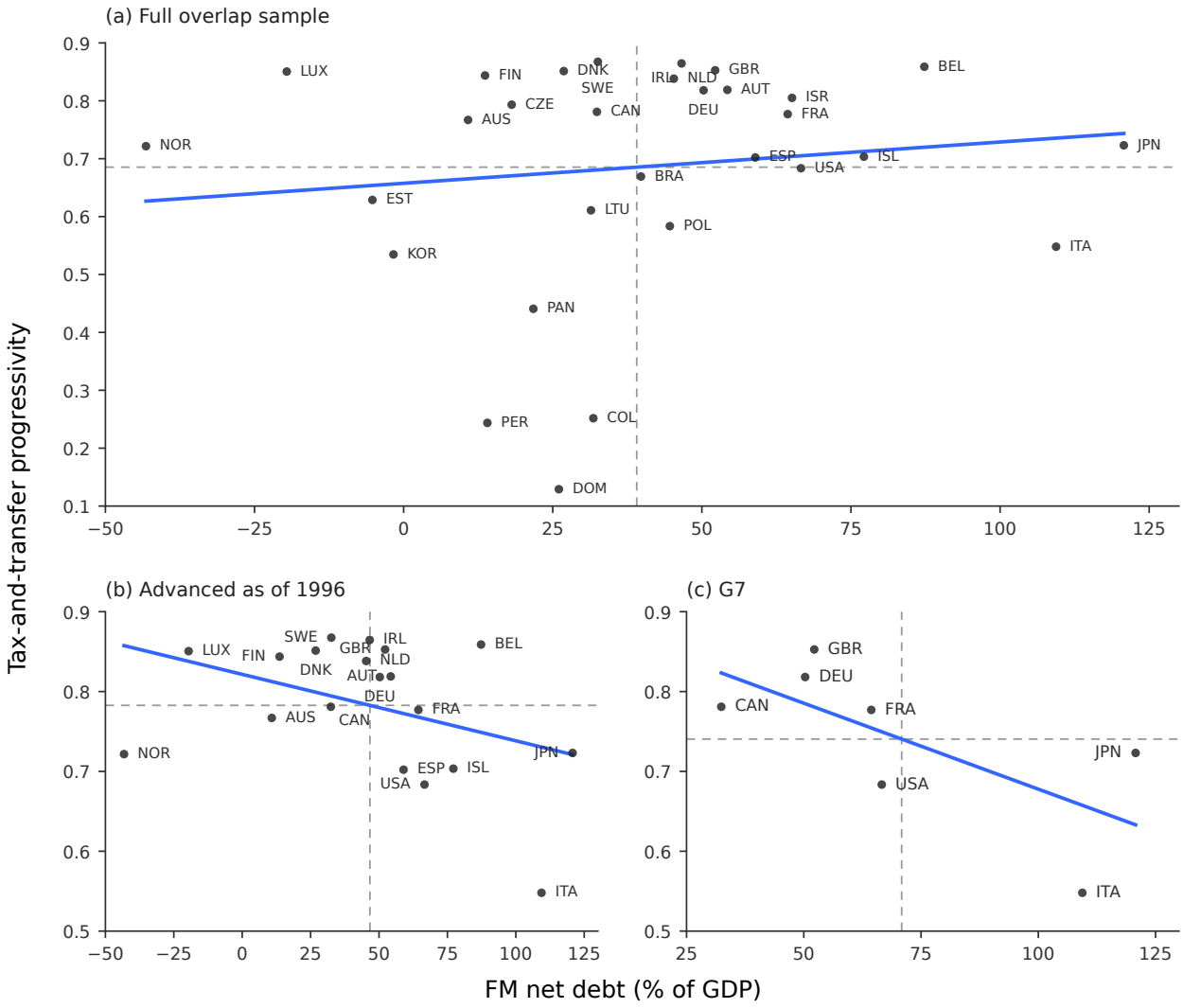
This appendix documents the cross-country patterns referenced in Section 4.4. Progressivity estimates come from **?**. I use two measures: one estimated from income taxes and another estimated from taxes and transfers. The latter is closest to the model, where p indexes the progressivity of taxes and transfers together. Debt comes from the IMF DataMapper series through 2019, using Global Debt Database (GDD) gross debt and Fiscal Monitor (FM) net debt. Each country enters as an average over the years where the progressivity estimates and the debt series overlap.[24]

Figure 21 shows how the relationship changes as the sample narrows, using the tax-and-transfer measure and net debt. In the full overlap sample, the rank correlation is 0.18 ($n = 30$). With gross debt, which has broader coverage, it is 0.35 ($n = 34$); after residualizing progressivity and debt on log GDP per capita, it falls to 0.23. Among countries classified as advanced by 1996, the correlation with net debt turns negative, at -0.35 ($n = 19$), and falls to -0.75 within the G7 ($n = 7$). The results with gross debt are nearly identical. The income-tax measure moves in the same direction: its correlation with net debt falls from 0.41 in the full sample to 0.09 among countries advanced by 1996 and -0.54 within the G7. Given the small samples, these patterns are descriptive.

Government size is another possible confound. Residualizing progressivity and net debt on government consumption leaves the full-sample correlation close to zero and the correlations in the two narrower samples negative: -0.43 among countries advanced by 1996 and -0.68 within the G7. Using total government expenditure instead gives similar results. Figure 22 plots the residualized relationship for the two narrower samples. Since government spending is itself a policy outcome, these are descriptive adjustments rather than causal estimates.

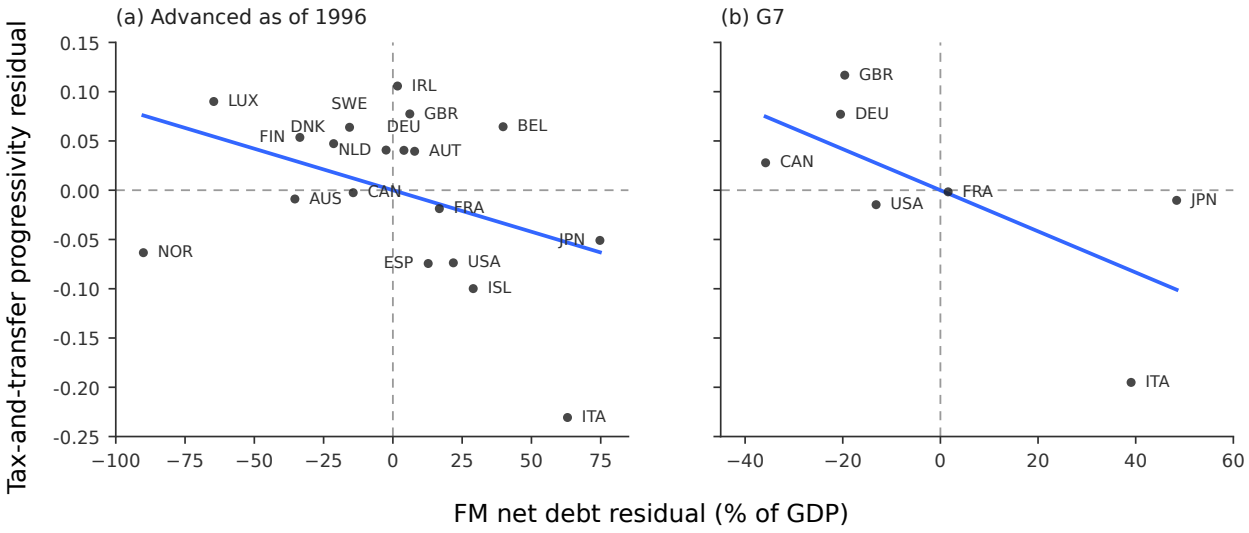
[24] The levels of the two progressivity measures are not directly comparable to the 0.181 used in the calibration, which comes from a different sample and income concept (**?**). These estimates are used only to compare progressivity across countries.

Figure 21: Progressivity and net debt as the sample narrows, 1980–2019



Note: Country means over matched years. Progressivity is the tax-and-transfer measure from **?**; debt is FM net debt. Panels (b) and (c) restrict the sample to economies advanced by 1996 and the G7. Axis ranges differ across panels. Solid lines are OLS fits.

Figure 22: Progressivity and net debt after controlling for government consumption



Note: Progressivity and net debt are residualized on general government final consumption within each sample. Horizontal scales differ across panels. Solid lines are OLS fits.