Taxes, debts, and redistributions with aggregate shocks*

Anmol Bhandari
bhandari@umn.edu

David Evans
dgevans@nyu.edu

Mikhail Golosov
golosov@princeton.edu

Thomas J. Sargent
thomas.sargent@nyu.edu

April 2015

Abstract

This paper models how transfers, a tax rate on labor income, and the distribution of government debt should respond to aggregate shocks when markets are incomplete. A planner sets a lump sum transfer and a linear tax on labor income in an economy with heterogeneous agents, aggregate uncertainty, and a single asset with a possibly risky payoff. Limits to redistribution coming from incomplete tax instruments and limits to hedging coming from incomplete asset markets affect optimal policies. Two forces shape long-run outcomes: the planner's desire to minimize the welfare cost of fluctuating transfers, which calls for a negative correlation between agents' assets and their skills; and the planner's desire to use fluctuations in the return on the traded asset to compensate for missing state-contingent securities. In a multi-agent model calibrated to match facts about US booms and recessions, the planner's preferences about distribution make policies over business cycle frequencies differ markedly from Ramsey plans for representative agent models.

KEY WORDS: Distorting taxes. Transfers. Redistribution. Government debt. Interest rate risk.
JEL CODES: E62,H21,H63

*We thank Mark Aguiar, Stefania Albanesi, Manuel Amador, Andrew Atkeson, Marco Bassetto, V.V. Chari, Harold L. Cole, Fatih Guvenen, Guy Laroque, Francesco Lippi, Robert E. Lucas, Jr., Ali Shourideh, Pierre Yared, Ellen Mcgrattan, and seminar participants at Bocconi, Chicago, EIEF, the Federal Reserve Bank of Minneapolis, IES, Princeton, Stanford, the Federal Reserve Bank of Atlanta, UCL, Universidade CatÃŻlica,Wharton, Swiss Finance Institute, 2012 Minnesota macro conference, Monetary Policy Workshop NY Fed for helpful comments.

If, indeed, the debt were distributed in exact proportion to the taxes to be paid so that every one should pay out in taxes as much as he received in interest, it would cease to be a burden.possible, there would be [no] need of incurring the debt. For if a man has money to loan the Government, he certainly has money to pay the Government what he owes it. Simon Newcomb (1865, p.85)

1 Introduction

What are the welfare costs of public debt? What determines whether and how quickly a government should retire its debt? How should tax rates, transfers, and government debt respond to aggregate shocks? We study how answers to these questions depend on a government's eagerness and ability to redistribute and to hedge fluctuations in deficits.

We restrict tax collections to be affine functions of labor income. Agents differ in their productivities and wealth. They trade a single security whose payoff possibly depends on aggregate shocks. A Ramsey planner attaches a vector of Pareto weights to different types of agents' discounted utilities and adjusts the proportional labor tax rate, transfers, and asset purchases in response to aggregate shocks. A distribution of assets gives rise to flows of asset earnings across agents that depend on how returns on the asset comove with aggregate shocks. These flows require the government to adjust labor taxes and transfers to achieve its distributive and financing goals. Labor taxes distort labor supplies, but fluctuations in transfers also lower welfare. A decrease in transfers in response to adverse aggregate shocks affects agents who have low present values of earnings especially.

Section 2 sets out preferences and possibilities that define the economic environment. Section 2.1 describes a Ricardian property for our environment that we exploit in concisely formulating a Ramsey problem. An individual's net asset position equals his assets minus the assets held be a particular benchmark agent. The vector of all agents' net asset positions affects the set of allocations that can be implemented in competitive equilibria with affine taxes. The insight that net and not gross asset positions matter reduces the dimension of a state sufficient to describe a Ramsey plan recursively. It also provides a way to separate roles played by government debt and transfers despite the Ricardian property. We do this by setting the asset holdings of the least productive agent always to zero and than backing transfers out from the difference between this agent's consumption and his after-tax labor income.

A Ramsey plan induces an ergodic distribution of transfers, the labor tax rate, and the distribution of assets that is determined by interactions between a) the government's ability to hedge aggregate shocks by taking advantage of fluctuations in the return on the asset it trades; and b) the government's preferences about redistribution that affect how costly it is to use fluctuations in transfers to hedge those shocks. The analysis in sections 3, 4, and 5 shows that these interactions shape the ergodic distribution of government debt in the following ways. If equilibrium outcomes make the return on the asset *low* when the net-of interest government deficit is high, then the government will run up debt. But if the return on the asset is *high* when the net-of-interest government deficit is high, the government will accumulate assets. The long run variances of government assets and the tax rate are both lower and rates of convergence to the ergodic distribution are higher in economies in which the comovement between the net-of interest deficit and the return on the asset is bigger in absolute magnitude. Other things equal, governments that want more redistribution eventually hold fewer assets.

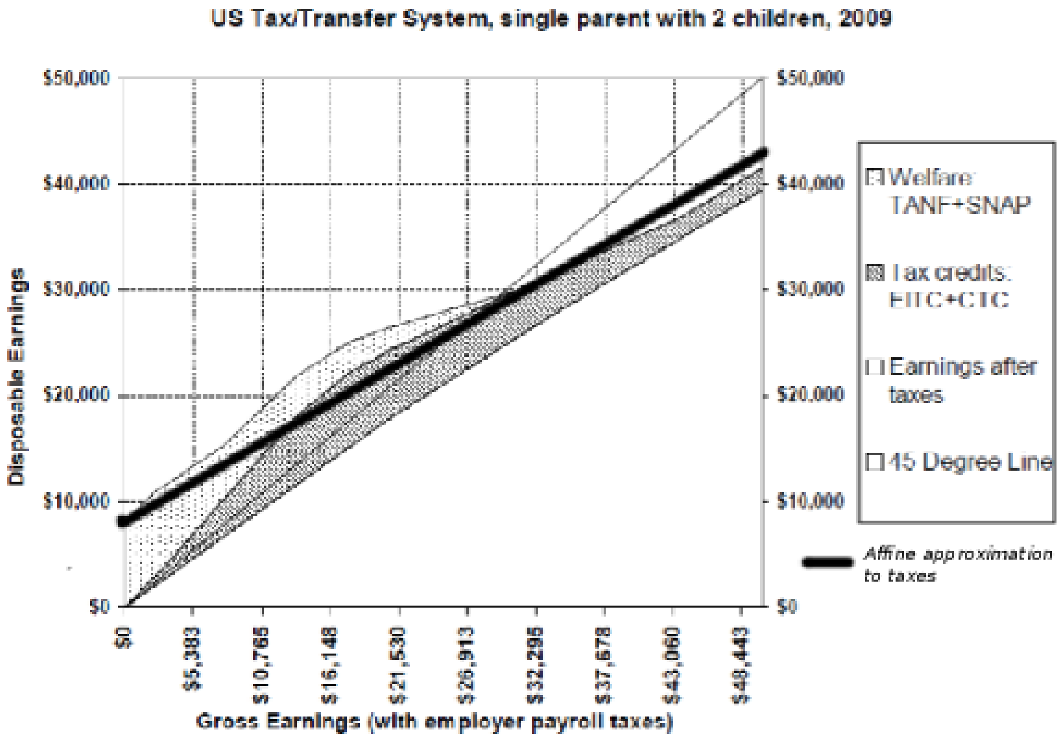
The comovement between the aggregate shock that drives government expenditures and returns on the single asset is a key intermediating object that shapes the ergodic joint distribution of the tax rate and government debt. In our general setting, one-period utilities are concave in consumption. That makes the return on the single asset an equilibrium object partly under the control of the Ramsey planner. To help us understand forces shaping both the ergodic joint distribution and rates of convergence to it, we begin by studying a special setting in which returns on the single asset and their correlation with the aggregate shock are exogenous, namely, a setting with one-period utilities that are quasilinear, meaning linear in consumption. Section 3 completes this analysis in two parts. We begin by analyzing a representative agent economy with quasilinear preferences, a government restricted to set transfers equal to zero always, and i.i.d aggregate shocks to public expenditures. Our main finding here is that for a large class of payoff structures, debt drifts towards an ergodic set that in a sense maximizes the government's ability to hedge expenditure shocks, a set that primarily depends on how the payoff on the asset correlate with fluctuations in the government's net-of-interest deficit. For special cases in which the asset payoff is affine in government expenditure shocks, we show that the ergodic distribution is degenerate. For other assumptions about the asset payoff, shock correlation, we develop tools to approximate the ergodic distribution and tell how the spread of the ergodic distribution of government debt and the tax rate increases with how far the payoffs are from allowing perfect fiscal hedging.

The representative agent analysis with transfers restricted to zero is informative about outcomes in multiple agent economies with no restrictions on transfers but with Pareto weights that make the welfare costs of transfers so high that the Ramsey planner chooses never to use them, or at least not to use them eventually.[1] We show this by analyzing an economy with unrestricted transfers and two types of agents both of whom have quasilinear one-period utilities and one of whom is not productive. We study the effects of the presence or absence of a nonnegativity constraint on the consumption of the unproductive agent as a determinant of the asymptotic level of government debt and whether and when transfers are used. The asymptotic level of assets is decreasing in the planner's desire for redistribution. This comes from the fact that welfare costs of using transfers are lower for a more redistributive government. Consequently it relies more on transfers and has less cause to accumulate assets to hedge aggregate shocks.

In section 4, we study economies more general in their heterogeneity, preferences, and shock structures. We formulate the Ramsey problem recursively in terms of two Bellman equation, one for time 0 and another for times $t \geq 1$. That these Bellman equations have different state variables expresses the time inconsistency of plans that as usual attribute to the Ramsey planner the ability to commit to an intertemporal plan at time 0. Subsections 4.2 and **??** derive conditions under which the planner can eventually achieve complete hedging, i.e., both constant labor tax rate and consumption shares even when the return on the asset is endogenous.

In section 5, we numerically verify that the forces isolated in the more analytically tractable models of sections 3 and 4 prevail in a version of the model with several types of agents calibrated to match US data. Our calibration captures (1) the initial heterogeneity wages and assets; (2) the observation that in recessions the left tail of the cross-section distribution of labor income falls by more than right tail; and (3) how inflation and asset return risk comove with labor productivity. We use this model to quantify the channels discovered in our theoretical analysis of simpler environments. We also describe features of optimal government policy, especially in booms and recessions at higher frequencies. During recessions accompanied by higher inequality,

[1]The analysis also augments what is known about representative agent economies in which a single risk-free bond is traded (e.g., Aiyagari et al. (2002), Farhi (2010), and Faraglia et al. (2012)).



it is optimal to increase taxes and transfers and to issue government debt. These outcomes differ both qualitatively and quantitatively from those in either a representative agent model or in a version of our model in which a recession is modelled as a pure TFP shock that leaves the distribution of skills unchanged.

Two technical contributions facilitate our analysis. First, for the quasilinear preferences studied in section 3, we obtain sharp characterizations of the ergodic distribution of debt and taxes by approximating transition dynamics around economies that permit perfect fiscal hedging. This approach also enables a local stability analysis for economies with the more general preferences of section **??**. Second, quantitative applications as ambitious as those in section 5 are made possible by taking a sequence of polynomial approximations about steady states of a sequence of deterministic economies and evaluating the polynomials at a current distribution of idiosyncratic state variables of the incomplete markets economy. This approach allows us to study transition paths of the optimal allocation that express the dynamics of the joint distribution of individual wealths and past consumptions, a high dimensional endogenous state vector. These technical contributions, which build on ideas developed in Evans (2014), will be useful in many other settings with heterogeneous agents and aggregate risk.

Section 6 offers concluding remarks.

1.1 Relationships to literatures

A large literature on Ramsey problems exogenously restricts transfers in the context of representative agent, general equilibrium models. Lucas and Stokey (1983), Chari et al. (1994), and Aiyagari et al. (2002) (henceforth called AMSS) Figure 1.1 shows that an affine structure better approximates the US tax-transfer system than just proportional labor taxes.

In contrast to those papers, our Ramsey planner cares about the distribution of welfare among agents with different skills and wealths. Except for not allowing them to depend on agents' personal identities, we leave transfers unrestricted and let the Ramsey planner set them optimally. We find that some of the same general principles that emerge from the representative agent, no-transfers literature continue to hold, in particular, the prescription to smooth distortions across time and states. However, it is also true that allowing the government to set transfers optimally changes the optimal policy in important respects.[2]

Some of our results in section 3 relate to Angeletos (2002), Buera and Nicolini (2004), and Shin (2007). These papers show that for a given initial government debt, the same Ramsey allocation that emerges from a complete markets like the one in Lucas and Stokey (1983) can also be supported in a with non-contingent debts with an appropriate set of maturities. In contrast, we characterize properties of payoffs on government debt that cause the optimal plan with incomplete markets eventually to converge to a complete markets allocation. That limiting allocation depends on model primitives but not on the initial level of government debt. We use this finding to build tools that allow us to say more about the ergodic distribution of debt and taxes in economies in which only imperfect fiscal hedging is possible.

In a setting with self-interested politicians who issue debt and tax capital and cannot commit to future government actions, Aguiar and Amador (2011) find conditions under which optimal allocations feature no tax distortions in the limit. Our results about conditions for eventual complete fiscal hedging in our environment have a similar flavor to Aguiar and Amador's. We delineate conditions under which fluctuations in the tax rates vanish in the limit, although the constant tax rate may not be zero. Also, for (a deterministic) setting with linear utility, Aguiar and Amador characterize speeds of convergence of debt and tax rates in neighbourhoods of steady states and they analyze how political economy frictions can arrest the rate of convergence. In section 3, we too characterize the speed of convergence and show how it depends on the covariance structure of the payoffs and aggregate shocks.

Our paper extends both Barro (1974), which showed Ricardian equivalence in a representative agent economy with lump sum taxes, and Barro (1979), which studied optimal taxation when lump sum taxes are ruled out. In our environment with incomplete markets and heterogeneous workers, both forces discovered by Barro play large roles. But the distributive motives that we include alter optimal policies.

Several other papers impute distributive concerns to a Ramsey planner. Three papers most closely related to ours are Bassetto (1999), Shin (2006), and Werning (2007). Like us, those authors allow heterogeneity and study distributional consequences of alternative tax and borrowing policies. Bassetto (1999) extends the Lucas and Stokey (1983) environment to include N types of agents with heterogeneous time-invariant labor productivities. There are complete markets. The Ramsey planner has access only to proportional taxes on labor income and state-contingent borrowing. Bassetto studies how the Ramsey planner's vector of Pareto weights influences how he responds to government expenditures and other shocks by adjusting the proportional labor

[2]A distinct strand of literature focuses on optimal policy in settings with heterogeneous agents when a government can impose arbitrary taxes subject only to explicit informational constraints (see Golosov et al. (2007) for a review). A striking result from that literature is that when agent's asset holdings are perfectly observable, the distribution of assets among agents is irrelevant and an optimal allocation can be achieved purely through taxation (see, e.g. Bassetto and Kocherlakota (2004)). In the previous version of the paper we showed that a mechanism design version of the model with unobservable assets generates some of the similar predictions to the model with affine taxes that we study, in particular, the relevance of net assets and history dependence of taxes. We leave further analysis along this direction to the future.

tax and government borrowing to cover expenses while manipulating competitive equilibrium prices to redistribute wealth between 'rentiers' (who have low productivities and whose main income is from their asset holdings) and 'workers' (who have high productivities) whose main income source is their labor.

Shin (2006) extends the AMSS (Aiyagari et al. (2002)) incomplete markets economy to two risk-averse households who face idiosyncratic income risk. When idiosyncratic income risk is big enough relative to government expenditure risk, the Ramsey planner chooses to issue debt so that households can engage in precautionary saving, thereby overturning the AMSS result that a Ramsey planner eventually sets taxes to zero and lives off its earnings from assets thereafter. Shin emphasizes that the government does this at the cost of imposing tax distortions. Constrained to use proportional labor income taxes and nonnegative transfers, Shin's Ramsey planner balances two competing self-insurance motives: aggregate tax smoothing and individual consumption smoothing.

Werning (2007) studies a complete markets economy with heterogeneous agents and transfers that are unrestricted in sign. He obtains counterparts to our Ricardian results about net versus gross asset positions, including the legitimacy of a normalization allowing government assets to be set to zero in all periods. Because he allows unrestricted taxation of initial assets, the initial distribution of assets plays no role. Our corollary 1 and corollary **??** generalize Werning's results by showing that all allocations of assets among agents and the government that imply the same net asset position lead to the same optimal allocation, a conclusion that holds for market structures beyond the complete markets structure analyzed by Werning. Werning (2007) provides an extensive characterization of optimal allocations and distortions in complete market economies, while we focus on precautionary savings motives for private agents and the government that are absent when markets are complete.[3,4]

Finally, our numerical analysis in Section 5 is related to McKay and Reis (2013). While our focus differs from theirs – McKay and Reis study the effect of a calibrated version of the US tax and transfer system on stabilization of output, while we focus on optimal policy in a simpler economy – both papers confirm the importance of transfers and redistribution over business-cycle frequencies.

2 Environment

We consider an infinite horizon economy in discrete time. There is a mass n_i of type $i \in I$ infinitely lived agents with $\sum_{i=1}^{I} n_i = 1$. Preferences of an agent of type i over stochastic processes for consumption $\{c_{i,t}\}_t$ and labor supply $\{l_{i,t}\}_t$ are represented by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U^i \left(c_{i,t}, l_{i,t}\right), \tag{1}$$

where $\mathbb{E}_t$ is a mathematical expectations operator conditioned on time t information, $\beta \in (0,1)$ is a time discount factor, $c_{i,t}$ and $l_{i,t}$ are consumption and labor supply of type i in period t, and U^i is the per-period utility function of type i.

[3] Werning (2012) studies optimal taxation with incomplete markets and explores conditions under which optimal taxes depend only on the aggregate state.

[4] More recent closely related papers are Azzimonti et al. (2008a,b) and Correia (2010). While these authors study optimal policy in economies in which agents are heterogeneous in skills and initial assets, they do not allow aggregate shocks.

Exogenous fundamentals include a stochastic cross section of skills $\{\theta_{i,t}\}_{i,t}$, government expenditures g_t, and the payoff p_t on a single asset traded between the government and the private sector. These are all functions of a shock $s_t \in S$ governed by an irreducible Markov process that takes values in a finite set S. We let $s^t = (s_0, ..., s_t)$ denote a history of shocks having joint probability density $\pi(s^t)$. To simplify exposition, we often use notation z_t to denote a random variable z with a time t conditional distribution that is a function of the history s^t. Occasionally, we use the more explicit notation $z\left(s^t\right)$ to denote a realization at a particular history s^t.

An agent of type i who supplies l_i units of labor in state s_t produces $\theta_i\left(s_t\right) l_i$ units of output, where $\theta_i(s_t) \in \Theta$ is a non -negative state-dependent scalar. Feasible allocations satisfy

$$\sum_{i=1}^{I} n_i c_{i,t} + g_t = \sum_{i=1}^{I} n_i \theta_{i,t} l_{i,t}. \tag{2}$$

The time t payoff p_t on the traded asset is described by

$$p_t = \mathbb{P}(s_t | s_{t-1}),$$

where $\mathbb{P}$ is an $S \times S$ matrix normalized to satisfy $\mathbb{E}_t p_{t+1} = 1$. Specifying the asset payoff in this way lets us investigate consequences of the correlation between asset returns, on the one hand, and government expenditures or shocks to the skill distribution, on the other hand. Purchasing $\check{b}_t$ units of the asset at a price of q_t units of time t consumption per unit of the asset entitles the owner to $p_{t+1}\check{b}_t$ units of time $t+1$ consumption. Consequently, $R_{t+1} = p_{t+1}/q_t$ is the gross rate of return on the asset measured in units of time $t+1$ consumption good per unit of time t consumption good. We let $b_t \equiv q_t \check{b}_t$ denote a time t value of $\check{b}_t$ units of the asset, measured in units of time t consumption. From now on, we express budget sets in terms of the gross rate of return R_t and counterparts of the value of assets b_t. To facilitate a unified notation for budget constraints for dates $t \geq 0$, we define $R_0 \equiv p_0 \beta^{-1}$.

Households and the government begin with assets $\{b_{i,-1}\}_{i=1}^{I}$ and B_{-1}, respectively. Asset holdings satisfy the market clearing condition

$$\sum_{i=1}^{I} n_i b_{i,t} + B_t = 0 \text{ for all } t \geq -1. \tag{3}$$

There is a proportional labor tax rate τ_t and common lump transfer T_t. The tax bill of an agent with wage earnings $l_{i,t}\theta_{i,t}$ is

$$-T_t + \tau_t \theta_{i,t} l_{i,t}.$$

A type i agent's budget constraint at $t \geq 0$ is

$$c_{i,t} + b_{i,t} = \left(1 - \tau_t\right) \theta_{i,t} l_{i,t} + R_t b_{i,t-1} + T_t, \tag{4}$$

and the government budget constraint is

$$g_t + B_t = \tau_t \sum_{i=1}^{I} n_i \theta_{i,t} l_{i,t} + R_t B_{t-1} - T_t. \tag{5}$$

Definition 1. An *allocation* is a sequence $\{c_{i,t}, l_{i,t}\}_{i,t}$. An *asset profile* is a sequence $\{\{b_{i,t}\}_i, B_t\}_t$. A *returns process* is a sequence $\{R_t\}_t$. A *tax policy* is a sequence $\{\tau_t, T_t\}_t$.

We impose debt limits on asset profiles. In general they can depend on the history of exogenous shocks s^t and the tax policy of the government $\{\tau_t, T_t\}_t$ and are described by functions $\{\underline{b}_i\left(s^t;\{\tau_t,T_t\}_t\right)\}_i$ and $\underline{B}\left(s^t;\{\tau_t,T_t\}_t\right)$. This allows both for natural debt limits, which correspond to the maximum debt that an agent can repay almost surely, and for tighter, ad-hoc limits. So for all $t \geq -1$ we impose,

$$b_{i,t} \geq \underline{b}_i(s^t;\{\tau_t,T_t\}_t) \tag{6}$$

and

$$B_{i,t} \geq \underline{B}(s^t;\{\tau_t,T_t\}_t). \tag{7}$$

Definition 2. For a given initial asset distribution $\left(\{b_{i,-1}\}_i, B_{-1}\right)$, a competitive equilibrium with affine taxes is a sequence $\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, B_t, R_t\}_t$ and a tax policy $\{\tau_t, T_t\}_t$ such that (i) $\{c_{i,t}, l_{i,t}, b_{i,t}\}_{i,t}$ maximize (1) subject to (4) and (6); and (ii) constraints (2), (3), (5), and (7) are satisfied.

A Ramsey planner's preferences over competitive equilibrium allocations are ordered by

$$\mathbb{E}_0 \sum_{i=1}^{I} \omega_i \sum_{t=0}^{\infty} \beta^t U_t^i\left(c_{i,t}, l_{i,t}\right), \tag{8}$$

where the Pareto weights satisfy $\omega_i \geq 0$, $\sum_{i=1}^{I} \omega_i = 1$.

Definition 3. Given $(\{b_{i,-1}\}_i, B_{-1})$, an optimal competitive equilibrium with affine taxes is a competitive equilibrium with an allocation that maximizes (8).

2.1 Relevant and irrelevant aspects of the asset distribution

Our economy features both distortionary taxation (since agent-specific lump sum taxes are not available) and debt limits. Despite these frictions, there is a sense in which the value of the government debt *by itself* provides little information about welfare costs of servicing it. The important statistic is not the level of government debt per se, but who holds it. To formalize this point, we define agents' *net assets* in period t as agents' asset holdings relative to those of some benchmark agent, $\{b_{i,t} - b_{1,t}\}_{i>1}$. The next proposition shows that it is the initial distribution of the net assets that determines welfare in the optimal competitive equilibrium.

Proposition 1. *For any pair of initial distributions* $\left(\{b'_{i,-1}\}_i, B'_{-1}\right)$ *and* $\left(\{b''_{i,-1}\}_i, B''_{-1}\right)$ *that satisfy the debt limits and*

$$b'_{i,-1} - b'_{1,-1} = b''_i - b''_{1,-1} \tag{9}$$

the values of (8) *at the optimal allocations are the same.*

We relegate the proof to the appendix 7.1. The intuition confirms the quote of Newcomb (1865) with which we began our paper. To see this, imagine an increase an initial level of government debt from 0 to some arbitrary level $B'_{-1} < 0$ when agents asset holdings are equal across agents. The optimal way to finance this increase in public debt is to reduce transfers and keep distortionary taxes $\{\tau_t\}_t$ unchanged. If asset holdings were equal to begin with, each unit of debt repayment achieves the same redistribution as one unit of transfers. The situation would be different if, say, richer people initially own disproportionately more government debt. A government with Pareto weights $\{\omega_i\}_i$ that favor equality would want to increase both transfers $\{T_t\}_t$ and the distorting labor tax rate $\{\tau_t\}_t$ to offset the increase in inequality associated with the increase in government debt.[5]

Proposition 1 holds for arbitrary debt limits. In general, the absolute levels of government debt will be determinate. As long as there are agents who face debt limits that are binding along the equilibrium path, changes in government debt can relax (or tighten) borrowing constraints and alter welfare.[6] As a way of contrast, in Corollary 1 we show that with natural debt limits, only net asset positions are meaningfully determined in any competitive equilibrium.

Corollary 1. *Given $\left(\{b_{i,-1}\}_i, B_{-1}\right)$, let $\left\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, B_t, R_t\right\}_t$ and $\{\tau_t, T_t\}_t$ be a competitive equilibrium with natural debt limits. For bounded sequences $\left\{\hat{b}_{i,t}\right\}_{i,t\geq -1}$ that satisfy*

$$\hat{b}_{i,t} - \hat{b}_{1,t} = \tilde{b}_{i,t} \equiv b_{i,t} - b_{1,t} \text{ for all } t \geq -1, i \geq 2, \tag{10}$$

there exist sequences $\left\{\hat{T}_t\right\}_t$ and $\left\{\hat{B}_t\right\}_{t\geq -1}$ that satisfy (3) and that make $\left\{\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_i, \hat{B}_t, R_t\right\}_t$ and $\left\{\tau_t, \hat{T}_t\right\}_t$ constitute a competitive equilibrium given $\left(\left\{\hat{b}_{i,-1}\right\}_i, \hat{B}_{-1}\right)$.

In some parts of the discussion below we will find it useful to normalize assets of the least productive agents to zero. With such normalization the usual intuition that higher level of government debt leads to higher distortions and lower welfare is recovered, due to the fact that it effectively corresponds to higher asset inequality. This normalization is helpful to relate our findings to studies of government debt in representative agent models.

3 Quasilinear preferences

In our general section 2 setting, the endogeneity of the return on the single asset and the multiplicity I of types of agents makes it difficult to parse the forces that govern how transfers should be used to smooth tax rate distortions and to redistribute across people. In this section, we analyze a special case that helps us isolate key forces. We assume two types of agents who have one-period utility functions that are quasi-linear in consumption. This makes the return on the asset exogenous and allows us to focus on how the comovement between returns and aggregate risk affects how transfers are used to smooth tax distortions and to redistribute.

Assumption 1. *Quasilinear preferences: $u(c,l) = c - \frac{l^{1+\gamma}}{1+\gamma}$.*

Assumption 2. $\theta_1 > \theta_2 = 0$.

Assumption 3. *IID shocks to expenditure: $g(s_t)$ is i.i.d over time.*

[5] This logic implies that government debt that is widely and evenly distributed (e.g., implicit Social Security debt) is less distorting than government debt owned mostly by people whose incomes are at the top of the income distribution (e.g., government debt held by hedge funds). It is possible to extend our analysis to open economy with foreign holdings of domestic debt. The more government debt is owned by the foreigners, the higher are the distorting taxes that the government needs to impose.

[6] This intuition is similar to that in Woodford (1990). The government can relax the ad-hoc debt limits because of the implicit assumption that it is easier to collect taxes than to enforce debt payments by citizens.

Assuming one type of agent who cannot work makes redistribution between rich and poor transparent, and i.i.d shocks help with analytical tractability. Without loss of generality we can restrict our attention to payoff matrices $\mathbb{P}$ that have identical rows, denoted by a vector P of dimension S such that $\sum_{s\in S}\pi(s)P(s) = 1$. At interior solutions, the first-order conditions for the household's problem for $l_{1,t}$ and $b_{i,t}$ imply that

$$\begin{aligned} l_{1,t}^{\gamma} &= (1-\tau_t)\theta, \\ R_t &= \frac{p_t}{\beta}, \end{aligned}$$

where we set $R_0 = \beta^{-1}p_t$. Substituting these into the budget constraint (4) gives

$$c_{1,t} + b_{1,t} = l_{1,t}^{1+\gamma} + \frac{p_t}{\beta}b_{1,t-1} + T_t, \quad t \geq 0, \tag{11a}$$

$$c_{2,t} + b_{2,t} = \frac{p_t}{\beta}b_{2,t-1} + T_t, \quad t \geq 0. \tag{11b}$$

Subtract the budget constraint of the unproductive agent (11b) from (11a) to eliminate transfers T_t and get

$$c_{1,t} - c_{2,t} + b_{1,t} - b_{2,t} = l_{1,t}^{1+\gamma} + \frac{p_t}{\beta}\left(b_{1,t-1} - b_{2,t-1}\right), \quad t \geq 0. \tag{12}$$

Equation (12) presents a recursive version of the implementability constraints appropriate for the quasilinear setting. Note from equation (12) that only relative assets, $b_{1,t} - b_{2,t}$ matter, so we adopt a normalization where we set the assets $b_{2,t}$ of the unproductive agent always to zero. This makes the asset market clearing condition imply that $nb_{1,t} = -B_t$ and also implies that transfers $T_t = c_{2,t}$. Next we make

Assumption 4. $c_{2,t} \geq 0$.

Requiring that $c_{2t} \geq 0$ is an easy way to make transfers too costly to a Ramsey planner who cares too little about the unproductive agent. Since the only income of the unproductive agents comes from transfers, the non-negativity constraints ensures that such transfers are always weakly positive.

We impose bounds on the planner's asset positions B_t. Given our normalization $b_{2,t} = 0$, bounds on planner's asset are equivalently restrictions on net assets positions. Imposing these bounds turns out to be useful as they map to debt limits in a representative agent economy, a special case of our quasilinear economy as we shall show below. Households are assumed to operate under debt limits that are looser than those implied by the bounds on the planner's assets. Let $(\omega, n) \in [0,1] \times [0,1]$ be the Pareto weight and mass assigned to the productive type 1 agent. We can exploit the recursive nature of constraint (12) to write the recursive problem that characterizes the optimal policy as a function of the beginning of the period government assets $B\_$.

$$V(B\_) = \max_{\{c_1(s),c_2(s),l_1(s),B(s)\}_s} \sum_s \pi(s)\left\{\omega\left(c_1(s) - \frac{l_1(s)^{1+\gamma}}{1+\gamma}\right) + (1-\omega)c_2(s) + \beta V(B(s))\right\} \tag{13}$$

where the maximization is subject to

$$c_1(s) - c_2(s) - n^{-1}B(s) = l_1(s)^{1+\gamma} - n^{-1}\beta^{-1}P(s)B\_, \tag{14a}$$

$$nc_1(s) + (1-n)c_2(s) + g(s) = n\theta l_1(s), \tag{14b}$$

$$c_2(s) \geq 0, \tag{14c}$$

$$\underline{B} \leq B(s) \leq \overline{B}. \tag{14d}$$

As we shall see in later sections, the government has two tools to smooth aggregate shocks: use fluctuations in asset returns to obtain state-contingent payoff or adjust taxes and transfers in response to a shock. Typically it is optimal to use both instruments simultaneously. In the quasilinear economy analysis simplifies because in the long run the planner only uses one of these tools. Which tool is used depends on how much the government cares about the redistribution. We show that if government's concern for redistribution is sufficiently low, i.e. the Pareto weight assigned to the productive agent is sufficiently high, then the government uses only fluctuations in assets returns to hedge the aggregate risk. Conversely, with sufficiently high concerns for redistribution the government eventually uses only fluctuations in transfers to smooth aggregate shocks. The threshold Pareto weight is defined by $\bar{\omega} = n\left(\frac{1+\gamma}{\gamma}\right)$ and the two regions exists as long as $n\left(\frac{1+\gamma}{\gamma}\right) < 1$. If it is not satisfied, we only have the second region.

Pareto weights $\omega > \bar{\omega}$

We start our analysis by considering the problem of the government which has a sufficiently low concern for redistribution. The key intermediate step is the following lemma

Lemma 1. *Suppose that $n < \frac{\gamma}{1+\gamma}$. If $\omega > \bar{\omega}$ then $T_t = 0 \quad \forall t \geq 0$.*

The proof is relegated to appendix 7.2. For Pareto weights attached to the productive agent above $\bar{\omega}$, using transfers is too costly for the planner because positive transfers subsidize the unproductive type 2 agent whose welfare the planner values too little. This makes the Ramsey outcomes identical to those for an economy in which a Ramsey planner is restricted to use a linear tax schedule (he cannot use transfers) and in which there is a representative agent whose allocation equals the allocation to the productive agent in our economy. The characterization of this case is of independent interest as we can compare our results with the literature, Lucas and Stokey (1983) and Aiyagari et al. (2002), both of which studied a representative agent economy with linear taxes.

The key force at play here is how market incompleteness, as captured by the structure of payoff on the asset P, impedes tax-smoothing in a joint ergodic distribution for the tax rate and government debt. To organize analytical results in this section, it is useful to collect some P vectors that are perfectly correlated with expenditure shocks g in the following set

$$\mathcal{P}^* = \left\{ P : \, P(s) = 1 + \frac{\beta}{B^*}(g(s) - \mathbb{E}g) \; \forall s \text{for some } B^* \in [\overline{B}, \underline{B}] \right\}. \tag{15}$$

Payoffs in the set $\mathcal{P}^*$ have a property that it is feasible for the government to perfectly hedge fluctuations in net-of interest deficits by holding debt level B^*. Note that for each $P^* \in \mathcal{P}^*$ there is a unique $B^*(P^*)$ that satisfies conditions of equation (15). The relationship between any such pair of $(P^*, B^*(P^*))$ can be written as

$$B^*(P^*) = \beta \frac{\text{var}(g(s))}{\text{cov}(P^*(s), g(s))}. \tag{16}$$

Finally, observe that the set of payoffs $\mathcal{P}^*$ is non-generic: it has a measure 0 in the space of possible asset payoffs.

Proposition 2. *Suppose that* $n < \frac{\gamma}{1+\gamma}$. *If* $\omega > \bar{\omega}$, *the value function* $V(B\_)$ *in equation* (13) *is strictly concave and the behavior of government assets under a Ramsey plan is characterized as follows:*

1. If $P \in \mathcal{P}^*$ then government assets converge to a degenerate steady state

$$\lim_t B_t = B^*(P) \quad a.s \quad \forall B_{-1}.$$

   There is $\tau^*(P)$ such that $\lim_t \tau_t = \tau^*(P) \quad a.s \quad \forall B_{-1}$.

2. If $P(s) \notin \mathcal{P}^*$ there exists an invariant distribution of government assets with the property,

$$\forall \epsilon > 0, \quad \Pr\{B_t < \underline{B} + \epsilon \text{ and } B_t > \overline{B} - \epsilon \quad i.o\} = 1.$$

   There is a $\tau(B)$ such that the tax rate $\tau_t = \hat{\tau}(B_t)$ and $\hat{\tau}' < 0$.

Moreover, if $P(s) - P(s') > \frac{\beta}{\underline{B}}[g(s) - g(s')] \quad \forall s, s'$ and $\overline{B} - \underline{B}$ is sufficiently large, there exist B_1, B_2 such that

$$\mathbb{E}V'(B(s)) > V'(B\_) \qquad for\ B\_ > B_2,$$

$$\mathbb{E}V'(B(s)) < V'(B\_) \qquad for\ B\_ < B_1.$$

The first part of Proposition 2 shows that if there is any asset level which allows the government to hedge its risk, it converges to that level almost surely. In that steady state the tax level is constant and the government uses fluctuations in asset return to finance stochastic expenditures. This dynamics resembles that of a complete market economy a-la Lucas and Stokey (1983), except the long-run level of debt and taxes in pinned down by vector P in our incomplete market economy, while with complete market economy it is pinned down by the initial value of government assets. Expression (16) tells us that whether the government eventually holds assets or owes debt is determined by the sign of the covariance of $P(s)$ with $g(s)$. In particular, the government accumulates positive (negative) assets if returns are high (low) in the states when $g(s)$ is high.

As we discussed above, the payoffs that allow the government to hedge its risks, are not generic. The second part of Proposition 2 shows that in general, when perfect hedging is impossible, the support of the invariant distribution is wide in the sense that almost all asset sequences recurrently revisit small neighborhoods of any arbitrary lower and upper bounds on government assets. With incomplete markets a long enough sequence of high or low shocks takes government assets arbitrarily far from any starting level of debt. Because the labor tax rate is decreasing in government assets, it varies too. These outcomes contrast sharply with those in a corresponding complete market benchmark like Lucas and Stokey's, where both debt and tax rates would be constant sequences, and with those in the incomplete markets economy of Aiyagari et al., where government assets would approach levels that allow the limiting tax rate to be zero and the tail allocation to be first-best. Finally, the last part of the proposition shows that with incomplete markets government assets exhibit a form of mean-revision: if the government accumulated sufficiently high or low level of assets, future debt levels revert towards the middle.

To acquire more information about the invariant distribution of government assets when the payoff vector P does not allow for perfect hedging, we consider linear approximations of the policy rules. For any P consider the closest perfect hedging payoff P^* that solves

$$\min_{P^* \in \mathcal{P}^*} \sum_s \pi(s) \left[P(s) - P^*(s)\right]^2 . \tag{17}$$

Use this P^* to find a component $\hat{P}$ of P that is orthogonal to g:

$$P(s) = \hat{P}(s) + P^*(s).$$

We linearize policy rules around $\hat{P} = \mathbf{0}$ and study the properties of the ergodic distribution generated by such rules.

Proposition 3. *The linearized policy rules induce a unique ergodic distribution of government debt with the following properties.*

- **Mean:** The ergodic mean of asset distribution, $\mathbb{E}(B)$, satisfies

$$\mathbb{E}(B) = B^*(P^*).$$

- **Variance:** The ergodic coefficient of variation of government assets B is

$$\frac{\sigma(B)}{\mathbb{E}(B)} = \sqrt{\frac{\text{var}(P(s)) - |\text{cov}(g(s), P(s))|}{(1 + |\text{cov}(g(s), P(s))|)|\text{cov}(g(s), P(s))|}} \leq \sqrt{\frac{\text{var}(\hat{P}(s))}{\text{var}(P^*(s))}}.$$

- **Convergence rate:** The rate of convergence to of the mean to its ergodic value is described by,

$$\mathbb{E}_{t-1}(B_t - B^*) = \frac{1}{1 + \text{var}(P)\text{corr}^2(P, g)}(B_{t-1} - B^*),$$

where $\text{corr}(P, g)$ is the correlation coefficient between P and g

We relegate the proof to appendix 7.4, where we describe how we take a first-order Taylor approximation to the decision rules and laws of motion for the state variables of our economy around complete market counterparts associated with $P^* \in \mathcal{P}^*$.

Proposition 3 describes how deviations from $\mathcal{P}^*$ map into larger variances for government debt and the tax rate under the ergodic distribution. The last part of Proposition 3 decomposes the the speed of convergence of assets B_t into two components: variance of $P(s)$ and the correlation between $P(s)$ and $g(s)$. If P were to belong to $\mathcal{P}^*$, the second term relating to correlation would be equal to one, and this bounds the (exponential) rate of convergence of the conditional mean (for payoff $P(s)$ close to $\mathcal{P}^*$) by variance of $P(s)$. Figure 1 illustrates how the ergodic distribution of government debt and the tax rate spread as we exogenously alter the covariance of $P(s)$ with $g(s)$.[7]

[7]For the purposes of these graphs we used the global approximation to the policy rules using standard projection methods rather than the linear approximations appealed to in Proposition 3. The ergodic distribution associated with the approximate linear policies is similar.

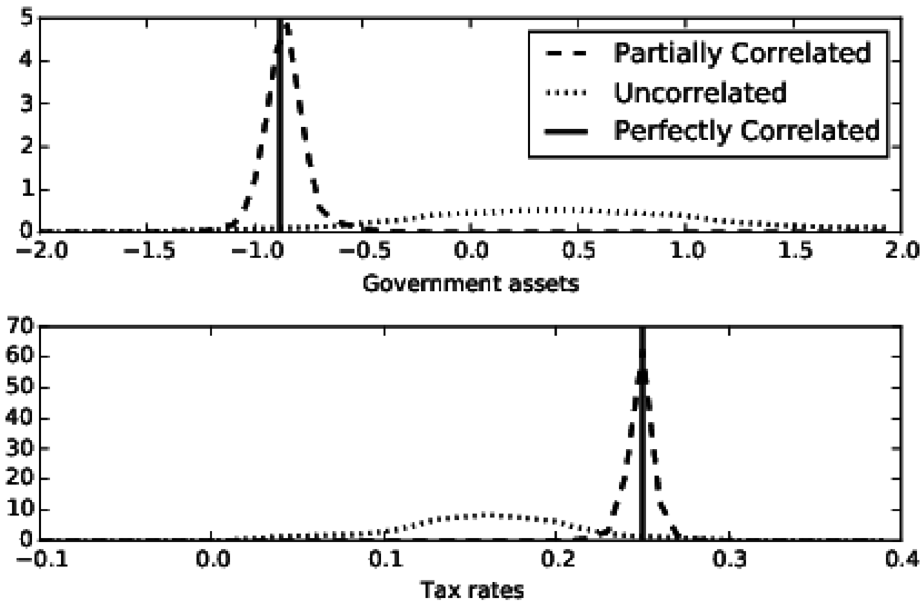


Figure 1: Ergodic distribution for government assets B_t and the labor tax rate τ_t in the representative agent quasilinear economy for three different asset payoff vectors P.

Pareto weights $\omega \leq \bar{\omega}$

The previous subsection described the optimal plan in the case where the Pareto weight on the productive type 1 agent was so high that the Ramsey planner chose not to use transfers to hedge aggregate shocks. Proposition 4 now completes the characterization by describing the case when the Ramsey planner has larger redistributive concerns and the Pareto weight attached to the productive agent are lower. This induces the planner to use transfers to hedge shocks.

Proposition 4. *For* $\omega < \bar{\omega}$*, and* $\min_s\{P(s)\} > \beta$*, there exist a* $\mathcal{B}(\omega)$ *satisfying* $\mathcal{B}'(\omega) > 0$ *and the optimal tax rate, transfer, and government asset policies* $\{\tau_t, T_t, B_t\}_t$ *are characterized as follows:*

1. If $B_{-1} > \mathcal{B}(\omega)$ then

$$T_t > 0, \quad \tau_t = \tau^*(\omega), \text{ and } B_t = B_{-1} \quad \forall t \geq 0.$$

2. If $B_{-1} \leq \mathcal{B}(\omega)$ then

   (a) If $P \in \mathcal{P}^*$ or $P \in \mathcal{P}^*$ and $B^*(P) > \mathcal{B}(\omega)$

$$T_t > 0 \text{ i.o.}, \quad \lim_t \tau_t = \tau^*(\omega) \text{ and } \lim_t B_t = \mathcal{B}(\omega) \quad a.s.$$

   (b) Otherwise,

$$\Pr\{\lim_t T_t = 0, \lim_t \tau_t = \tau^{**}, \lim_t B_t = B^*(P) \text{ or } T_t > 0 \text{ i.o. and} \lim_t \tau_t = \tau^*(\omega), \lim_t B_t = \mathcal{B}(\omega)\} = 1.$$

Greater concerns for redistribution lower costs of transfers and this expands the planner's hedging possibilities against aggregate shocks. Proposition 4 asserts that the ergodic distribution is degenerate and long run level of assets in general depends on both the initial assets and how much the planner values the welfare of the unproductive agent. If the government begins with enough assets (as in part 1 of Proposition 4), the planner chooses an interior allocation in which *all* fluctuations in net-of interest deficits are always financed by fluctuating transfers T_t. Here government assets remain at their initial level and the tax rate is constant at a level that is independent of the initial assets.

Part 2 describes a setting in which government assets are initially too low to validate the part 1 outcome. Part 2a shows that if perfect spanning is not feasible, then the planner eventually accumulates government assets until they reach the threshold $\mathcal{B}(\omega)$ at which the welfare costs of transfers are low enough to induce the planner to keep the tax rate constant.[8] Pareto weights that indicate that the planner cares more about the unproductive agent lower the marginal welfare costs of collecting revenue from labor taxes paid by the productive agent. This induces the planner to increase the labor tax rate, lowering the threshold level of assets that are required to finance all shocks by transfers. In this way, a more redistributive planner relies more on transfers and consequently is less motivated to accumulate assets to hedge aggregate shocks. Part 2b shows that if $P \in \mathcal{P}^*$, then assets converge either to the bound $\mathcal{B}(\omega)$ or to the perfect spanning asset level $B^*(P)$.

4 Optimal allocation with risk aversion

In this section we extend our analysis to economies with risk aversion. The main difference from the analysis in the previous section is that the return $R(s)$ depends both on the structure of asset payoffs and on the marginal utilities of consumption of agents. The latter term in general, depends on the realization of shocks and government policy. We begin our analysis with a pair of Bellman equations that formulate the Ramsey plan recursively for the general settings with $I > 2$ agents and one-period utility functions that are concave in consumption. We assume that $U^i : \mathbb{R}^2 \to \mathbb{R}$ is concave in $(c, -l)$ and twice continuously differentiable. We let $U^i_{x,t}$ or $U^i_{xy,t}$ denote first and second derivatives of U^i with respect to $x, y \in \{c, l\}$ in period t. Finally, we impose natural debt limits, which ensures that all first order conditions are interior.

In this section, we restrict attention to interior solutions and standard steps (see e.g. Chari et al. (1994)) ensure that $\left\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, R_t, \tau_t, T_t\right\}_t$ is a competitive equilibrium allocation if and only if they satisfy the constraints,

$$(1-\tau_t)\,\theta_{i,t}U^i_{c,t} = -U^i_{l,t}, \tag{18}$$

$$U^i_{c,t} = \beta \mathbb{E}_t R_{t+1} U^i_{c,t+1}, \tag{19}$$

$$c_{i,t} + b_{i,t} = -\frac{U^i_{l,t}}{U^i_{c,t}} l_{i,t} + T_t + \frac{p_t U^i_{c,t-1}}{\beta \mathbb{E}_{t-1} p_t U^i_{c,t}} b_{i,t-1} \;\; \forall i \geq 1, t \geq 1, \tag{20}$$

[8] The planner's choice under the conditions of part 2 to accumulate enough government assets so that eventually the government can use earnings on its assets together with fluctuating positive transfers to hedge fiscal shocks is reminiscent of outcomes in the Aiyagari et al. (2002) economy. There, with a representative agent and non-negativity constraints on transfers, the planner accumulated enough assets to finance shocks with zero distortionary labor taxes while costlessly using transfers to dispose of any excess earnings on its asset holdings. With multiple agents, fluctuating transfers may bring welfare costs that depend on the Ramsey's planner attitude about redistribution.

$$c_{i,0}+b_{i,0}=-\frac{U^i_{l,0}}{U^i_{c,0}}l_{i,0}+T_0+p_0\beta^{-1}b_{i,-1}\ \forall i\geq 1, \tag{21}$$

together with feasibility (2). We can simplify (20) and (21) by substituting for T_t and deriving the implementability constraints in terms of net asset positions as in of equation (12):

$$\begin{aligned}&(c_{i,t}-c_{I,t})+\tilde{b}_{i,t}\\ =\ &-\frac{U^i_{l,t}}{U^i_{c,t}}l_{i,t}+\frac{U^I_{l,t}}{U^I_{c,t}}l_{I,t}+\frac{p_tU^i_{c,t-1}}{\beta\mathbb{E}_{t-1}p_tU^i_{c,t}}\tilde{b}_{i,t-1}\text{ for } i<I \text{ and } t\geq 1.\end{aligned} \tag{22}$$

We are now ready to write the problem recursively. Let $\boldsymbol{x}=\left(U^I_c\tilde{b}_1,...,U^I_c\tilde{b}_{I-1}\right)$, $\boldsymbol{\rho}=\left(U^1_c/U^I_c,...,U^{I-1}_c/U^I_c\right)$, and $\mathcal{R}\left(s|s\_\right)\equiv P(s|s\_)\left(\frac{U^I_c(s)}{\mathbb{E}_{s\_}PU^I_C}\right)$. Following Kydland and Prescott (1980) and Farhi (2010), we split the Ramsey problem into a time-0 problem that takes $(\{\tilde{b}_{i,-1}\}_{i=1}^{I-1},s_0)$ as state variables and a time $t\geq 1$ continuation problem that takes $(\boldsymbol{x},\boldsymbol{\rho},s\_)$ as state variables. For $t\geq 1$, let $V(\boldsymbol{x},\boldsymbol{\rho},s\_)$ be the planner's continuation value given $\boldsymbol{x}_{t-1}=\boldsymbol{x},\boldsymbol{\rho}_{t-1}=\boldsymbol{\rho},s_{t-1}=s\_$. Let $a(s)=\{c_i(s),l_i(s)\}_i$ denote the allocation, the optimal allocation solves:

$$V(\boldsymbol{x},\boldsymbol{\rho},s\_)=\max_{\{a(s),x'(s),\rho'(s)\}}\sum_s\pi(s|s\_)\left(\left[\sum_i\omega_iU^i(s)\right]+\beta V(\boldsymbol{x}'(s),\boldsymbol{\rho}'(s),s)\right) \tag{23}$$

where the maximization is subject to

$$U^I_c(s)\left[c_i(s)-c_I(s)\right]+\left(\frac{U^i_l(s)l_i(s)}{U^i_c(s)}U^I_c(s)-l_I(s)U^I_l(s)\right)+x'_i(s)=\frac{x_i\mathcal{R}(s|s\_)}{\beta}\text{ for all } s,i<I \tag{24a}$$

$$\frac{\mathbb{E}_{s\_}P\boldsymbol{U}^i_c}{\mathbb{E}_{s\_}P\boldsymbol{U}^I_c}=\rho_i\text{ for all } i<I \tag{24b}$$

$$\frac{U^i_l(s)}{\theta_i(s)U^i_c(s)}=\frac{U^I_l(s)}{\theta_I(s)U^I_c(s)}\text{ for all } s,i<I \tag{24c}$$

$$\sum_i n_ic_i(s)+g(s)=\sum_i n_i\theta_i(s)l_i(s)\quad\forall s \tag{24d}$$

$$\rho'_i(s)=\frac{U^i_c(s)}{U^I_c(s)}\text{ for all } s,i<I \tag{24e}$$

Constraints (24b) and (24e) imply (19). The definition of x_t and constraints (24a) together imply equation (22) scaled by U^I_c. Period 0 maximization problem is similar, except that it does not have constraints (24b). For completeness we provide the time $t=0$ Bellman equation in appendix 7.6.

4.1 A formula for optimal taxes

We summarize the main tradeoff that the planner faces in setting labor taxes and transfers in the following proposition. Let $\{\mu_i(s)\}_{i<I},\xi(s)$ be the multipliers on constraints (24a) and (24d).

Proposition 5. *Suppose preferences are separable and have a constant Frisch elasticity, $U^i(c,l)=U^i(c)-\frac{l^{1+\gamma}}{1+\gamma}$. Then the optimal labor tax rate can be expressed as,*

$$\frac{1}{1-\tau(s)} = -\left(\frac{\bar{w}(s)\bar{y}(s) + cov\left(w_i(s), y_i(s)\right)}{\xi(s)\bar{y}(s)}\right), \tag{25}$$

where $w_i(s) = (1+\gamma)\hat{\mu}_i(s) + \gamma\left[n_i^{-1}\omega_i U_c^i(s)\right]$, $\bar{w}(s) = \sum_i n_i w_i(s)$, $\bar{y}(s) = \sum_i n_i y_i(s)$ and $\hat{\mu}_i(s) = U_c^I(s)\mu_i(s) - n_i^{-1}\omega_i U_c^i(s)$ for $i < I$ and $\hat{\mu}_I(s) = -\frac{\sum_{i<I}\hat{\mu}_i(s)n_i}{n_I} - n_I^{-1}\omega_I U_c^I(s)$.

The weights $\{w_i\}_{i=1}^I$, defined in this Proposition, show how the social planner evaluates fluctuations in inequality in this economy. It consists of Pareto-weight adjusted marginal utility of consumption of the agents adjusted to the cost of raising revenues, represented by the Lagrange multiplier $\{\hat{\mu}_i\}_{i=1}^I$.[9]

This equation shows the main trade-off that the social planner faces with heterogeneous agents. The planner can respond to an aggregate shock either by adjusting transfers, which leads to fluctuation in equality represented by $\left\{\omega_i U_c^i\right\}_{i=1}^I$, or by changing taxes, the deadweight loss of which is represented by $\{\hat{\mu}_i\}_{i=1}^I$. In the optimum the planner equilizes distortions from the two policies. All things being equal, labor taxes are high if inequality in consumption is high or the deadweight loss of taxation is low.

We discuss two scenarios where things simplify and we can interpret these weights more tightly. The first case is when preferences are quasilinear and our formula (25) reduces to

$$\frac{1}{1-\tau(s)} = 1 - \left(\frac{cov(n_i^{-1}\omega_i, y_i(s))}{\gamma^{-1}\bar{y}(s)}\right). \tag{26}$$

Since marginal utilities are constant, planner's concerns for equity are captured by the covaraince between exogenous Pareto weights and pre-tax income. The covaraince is negative when Pareto weights are higher on agents with low pre-tax labor income and this calls for higher labor taxes. The covariance and the optimal taxes vary across states if the distribution of skills $\{\theta_i(s)\}$ changes with aggregate shocks, otherwise we have constant tax rates.[10]

Formula (25) is closely related to classical results on optimal commodity taxation and redistribution derived Diamond (1975) and Atkinson and Stiglitz (1976). These papers studied optimal commodity taxes in static settings and derived formulae for optimal taxes similar expressed in terms of average elasticities on one hand and cross sectional covariances that captured equity considerations on the other. In absence of savings, our problem (23) is a special case of their setting and formula (25) reduces to

$$\frac{1}{1-\tau} = -\left(\frac{cov(n_i^{-1}\omega_i U_c^i, y_i) - \bar{y}((1+\gamma)\gamma^{-1}\xi - 1)}{\xi\bar{\varepsilon}}\right), \tag{27}$$

where $\sigma_i = -\frac{U_{cc}^i(s)c_i(s)}{U_c^i(s)}[1 - \frac{T(s)}{c_i(s)}]$ as the elasticity of substitution with respect to consumption (net of transfers) for agent i, and $\bar{\varepsilon} = \sum_i \left(\frac{1-\sigma_i}{\sigma_i+\gamma}\right) n_i y_i$.

[9] In particular $\hat{\mu}_i(s)$ is agent i's contribution to the planner's total cost of reducing agent i asset holdings and simultaneously increasing transfers by one unit.

[10] In section 3, we imposed non-negativty constraints on the consumption of the unproductive agent as a way of modeling costs of transfers that typically come from spreading of marginal utilities for more general preferences in adverse aggregate shocks. As summarized by Theorems (2)-(4), these can vary across states even if the skill distribution is unchanged.

4.2 Eventual complete hedging with binary shocks

In section 3 with quasilinear preferences, we studied scenarios where optimal policy perfectly hedged aggregate shocks. Although the conditions required were restrictive, specifically that the payoff vectors had to be perfectly correlated with expenditure shocks, these economies served as useful benchmark and a point of approximation to study optimal policy in more general cases where perfect hedging was not feasible for any asset levels. In this section we extensions of those insights to the economy with risk-aversion.

With risk aversion, perfect spanning corresponds to an allocation such that the (adjusted) payoffs $\mathcal{R}(s|s\_)$ offset movements in $U_c^I(s)\left[c_i(s) - c_I(s)\right] + \left(\frac{U_l^i(s)}{U_c^i(s)} U_c^I(s) - l_I(s) U_l^I(s)\right)$ for all agents keeping $\boldsymbol{x(s)} = \boldsymbol{x}$ and at the same time provide full insurance i.e $\boldsymbol{\rho}(s) = \boldsymbol{\rho}$. For a given state $(\boldsymbol{x}, \boldsymbol{\rho}, s\_)$, let $\Psi\left(s; \boldsymbol{x}, \boldsymbol{\rho}, s\_\right) = \left(x'\left(s\right), \rho'\left(s\right)\right)$ solve the value functions (23) so that $\Psi\left(s; \boldsymbol{x}, \boldsymbol{\rho}, s\_\right)$ is the law of motion for the state variables under a Ramsey plan at $t \geq 1$.

Definition 1. *A steady state satisfies* $\left(\boldsymbol{x}^{SS}, \boldsymbol{\rho}^{SS}\right) = \Psi\left(s; \boldsymbol{x}^{SS}, \boldsymbol{\rho}^{SS}, s_{-}\right)$ *for all* $s, s\_$.

At such *steady states*, outcomes resemble those in the complete market economy of Werning (2007). The tax rate and transfers both depend only on the current realization of shock s_t.[11] Arguments of Werning can be adapted to show that the tax rate is constant when preferences have the CES form $c^{1-\sigma}/(1-\sigma) - l^{1+\gamma}/(1-\gamma)$ and relative skills are constant across aggregate shocks. We discuss the conditions for existence of steady states and then study a simple example that will allow us to identify forces also present in the quasilinear economy of section 3 and the more general economies to be analyzed with numerical methods in section 5.

Lemma 2. *When utility is strictly concave in consumption,* $\|S\| = 2$ *is necessary for a steady state to exist generically.*

The existence of steady state depends on the solution to a non-linear system, which can be verified only numerically. Appendix 7.8 discusses the structure of these equations in more detail. Here we present an example with risk averse agents where existence and comparative statics can be established analytically.

Our example is an economy consisting of two types of households with $\theta_1 > \theta_2 = 0$ and common one-period utilities $\ln c - \frac{1}{2} l^2$. The shock s takes two values $\{s_L, s_H\}$ that are i.i.d across time. We assume that $g\left(s_H\right) > g\left(s_L\right)$. With natural debt limits, Corollary 1 applies and we are free to normalize the assets of the unproductive agent to zero and interpret $x(s) = U_c^I(s) B(s)$ as the marginal utility adjusted government's assets. In this example we show,

Proposition 6. *Suppose that* $g(s) < \theta_1$ *for all* s *and let* $\mathcal{R}(s|x,\rho) \equiv \frac{P(s) U_c^2(s|x,\rho)}{\mathbb{E} P U_c^2(x,\rho)}$ *be the adjusted payoff vector. If* $P\left(s\right) = 1$ *for all* s*, then there exists a steady state* (x, ρ) *with* $x > 0$ *and* $\mathcal{R}\left(s_L|x,\rho\right) < \mathcal{R}\left(s_H|x,\rho\right)$. *In addition, there exists a payoff vector* P *such that* $x < 0$ *and* $\mathcal{R}\left(s_L|x,\rho\right) > \mathcal{R}\left(s_H|x,\rho\right)$.

Proposition 6 establishes the existence of steady states and isolates how the comovement of the return on the asset with expenditure shocks predicts the sign of government assets.[12] In the quasilinear case we could disentangle the considerations for spanning and redistribution

[11] However note that the distribution of assets in the steady state for us is endogenous, it only depends on the technology and preference parameters, as against in Werning (2007) where it is pinned down by the initial distribution of assets.

[12] This parallels the quasilinear case, see expression (16)

explicitly. With risk aversion, costs of transfers come from fluctuations in marginal utilities and are interlinked with fluctuations in endogenous returns. A decrease in transfers in states with high expenditure disproportionately affects the low-skilled agent, so his marginal utility falls by more than the marginal utility of a high-skilled agent. Choosing policies such that the net assets of the high-skilled agent decrease over time makes the two agents' after-tax and after-interest income become closer, allowing decreases in transfers to have smaller effects on inequality in marginal utilities. This force pushes the government to accumulate assets in the long run. On the other hand, if (adjusted) payoffs are sufficiently low in states with high expenditures, holding negative assets or debt is optimal from hedging perspective; as doing so lowers the interest rate burden. The sign of long run assets ultimately depends on the balance of these forces.

5 Quantitative investigation

In sections 3 and 4, we studied steady states as a way of summarizing the asymptotic behavior of Ramsey allocations in particular distribution of assets and taxes. In this section, we use a calibrated version of the economy a) to revisit the magnitude of these forces; and b) to study optimal policy responses at business cycle frequencies when the economy is possibly far away from a (stochastic) steady state.

We calibrate parameters and shock process to reflect stylized facts about dynamics of U.S. distribution of earnings, assets, and returns on government assets for the period 1978-2010. We match those facts in a competitive equilibrium with an arbitrary government policy chosen to reflect the U.S. tax-debt policy for the same period. We then solve for the optimal Ramsey allocation for the same initial distribution of relative assets with Pareto weights chosen to match the average tax level in the competitive equilibrium. We can then compare outcomes predicted under the optimal policy to actual U.S. data. The numerical calculations use methods adapted from Evans (2014) and described in the appendix 7.10.

5.1 Calibration

To facilitate calibration, we construct a competitive equilibrium with incomplete asset markets using an arbitrary tax-debt policy that fits US tax, debt and expenditure data for the period 1978-2010. We set the tax and debt policy to follow

$$\log B_t = (1-\rho_B)\log \bar{B} + \rho_{B,B}\log B_{t-1} + \rho_{B,Y}\log Y_t + \sigma_B \epsilon_{B,t},$$

$$\tau_t = \bar{\tau} + \rho_{\tau,Y}\log Y_t + \sigma_\tau \epsilon_{\tau,t}.$$

To estimate the parameters that describe these rules we use (HP) filtered annual data on debt, aggregate labor earnings, and the time series for average marginal (federal) tax rates from Barro and Redlick (2009). Government expenditures are exogenous and follow

$$\log g_t = \log \bar{g} + \sigma_g \epsilon_{g,t}.$$

Transfers are obtained as a residual from the government budget constraint. Table 1 summarizes the estimates.

We use $I = 9$ so that our agents represent 10th to 90th quantiles of (working) US households labor earnings distribution. To calibrate the wealth inequality, we use Survey of Consumer Finances [1989-2007] to compute average net assets by earnings quantiles. We set initial assets such that the relative assets match the SCF and they add up to the total federal debt held by the public. We assume that all households have isoelastic preferences $u(c,l) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{l^{1+\gamma}}{1+\gamma}$. We set $\sigma = 2.5$ and choose $\gamma = 2$ to target a Frisch elasticity of labor supply of 0.5. We set the time discount factor $\beta = 0.98$, which implies the annual interest rate in an economy without shocks would be 2% per year.

Our theoretical analysis emphasized two forces that determine cyclicality of the optimal policy: the correlation of income inequality and asset returns with aggregate shocks. To parsimoniously capture the behavior of labor income inequality, we assume that agents productivity follows a two-parameter family of stochastic processes

$$\log \theta_{i,t} = \log \bar{\theta}_i + \epsilon_t[1 + (.9 - \mathcal{Q}(i))m], \tag{28}$$

where $\mathcal{Q}(i)$ is the percentile of agent i. In this specification ϵ_t is a common aggregate productivity shock, and parameter m governs relative volatility of wages of agents in different quantiles of labor income distribution. For example, when $m = 1./0.8 > 0$ the percentage decline in productivity in a recession for agents in the 10th percentile is twice that of the 90th percentile. Aggregate productivity shock followsan AR(1) process

$$\varepsilon_t = \rho_\varepsilon \varepsilon_{t-1} + \sigma_\varepsilon \epsilon_{\theta,t}.$$

As in the Ramsey taxation problem, the agents trade a single asset with exogenous payoffs p_t. We assume

$$p_t = 1 + \chi \epsilon_t,$$

where χ captures the ex-post comovement in returns on government assets and aggregate shocks. The parameters $\theta_i, m, \rho_\epsilon, \sigma_\epsilon$ and χ are calibrated so the a competitive equilibrium matches the following moments for the U.S. economy in the 1978-2010 period.

We set $\{\bar{\theta}_i\}$ such that the dispersion earnings in the competitive equilibrium matches the average dispersion per quantile as reported by Guvenen et al. (2014).[13] To calibrate m use use the finding in Guvenen et al. (2014) that in the past four recessions, the average fall in income for agents in the first decile of earnings was on an average three times that experienced by the 90th percentile. Furthermore, between the 10th and the 90th percentiles, the change in the percentage drop in earnings was almost linear. From these facts we infer a slope $m = 0.90$. The persistence and standard deviation of the productivity shock, $\rho_\epsilon, \sigma_\epsilon$ are set to match auto-correlation and the standard deviation of aggregate labor earnings. Finally the real return on 3 month treasury

[13] These authors use recently available high quality administrative data to document labor income loss across for different percentiles of earning distribution in the last four US recessions These include 1979-83, 1990-92, 2000-02, 2007-10 recessions.

Parameter	Value	Parameter	Value
$\bar{B}$	0.6	$\bar{\tau}$	0.24
$\rho_{B,B}, \rho_{B,Y}$	(0.87, -1.21)	$\rho_{\tau,Y}$	0.09
σ_B	0.023	σ_τ	0.006

Table 1: Estimates for US tax-debt policies

bill is acylical in the data for our sample. We set $\chi = 0.85$ to make the $\frac{p_t}{q_{t-1}}$, which is the return on government assets in our model approximately acylical.

Lastly Pareto weights are set from a one parameter family such that the average tax rate under the optimal allocation is equal to $\bar{\tau} = 0.24$. Table 2 summarizes the parameters underlying the benchmark calibration.

Parameter	Value	Description
γ	2	Average Frisch elasticity of labor supply of 0.5
β	0.98	Average (annual) risk free interest rate of 2%
σ	2.5	Standard
m	0.9	Relative 90-10 drop in recessions = 3.2
χ	0.85	Covariance between holding period returns and output
$\sigma_\epsilon, \rho_\epsilon$	0.015, 0.55	Standard deviation and autocorrelation of aggregate labor earnings
$\bar{g}, \sigma_g$	.10 %, 0.014	Federal government consumption expenditures

Table 2: Benchmark calibration

5.2 Long run outcomes

In this section we verify that the main findings in our theoretical analysis of the quasi linear settings hold in more general settings.

We being with Proposition 3 that points to the importance of correlation of payoffs and aggregate shocks for the speed and the long run average level of taxes and debt. In the more general case, as analyzed in this section, the correlation between marginal utility adjusted payoffs, $Corr[U_c P, \epsilon]$, will govern these properties. A sufficiently positive χ generates lower payoffs in recessions relative to booms which makes $Corr[UcP, \epsilon]$ less negative. Figure 2 simulations of 5000 periods for the government debt to output ratio, the labor tax rate, and the transfers to output ratio for three values of χ that generate $Corr[UcP, \epsilon] \in \{-0.88, -0.60(\text{benchmark}), -0.22\}$ in dashed, bold, and dotted, respectively. The three simulations start from the same initial conditions and all share the same sequence of underlying shocks.

Two features emerge. Different values χ give rise to different locations of the long-run marginal distribution of government assets and also to different rates of convergence to that long-run distribution. In line with assertions of Proposition 3, the long run level of assets is inversely related to the magnitude of this correlation. Also the speed of convergence is slower when the absolute magnitude of $Corr[UcP, \epsilon]$ is low.

In order to get a clearer picture of the speed of convergence, we plot paths of the conditional means for debt and the tax rate in figure 3. To explain how we generated these plots, let $B(s_{t+1}, \boldsymbol{x}_t, \boldsymbol{\rho}_t)$ be the Ramsey decision rules that generate the assets B of the government and let $\Psi\left(s_{t+1}; \boldsymbol{x}_t, \boldsymbol{\rho}_t\right)$ be the law of motion for the state variables for the Ramsey plan. For a given history, the conditional mean of government assets is:

$$B_{t+1}^{cm} = \mathbb{E}B(s_{t+1}, \boldsymbol{x}_t^{cm}, \boldsymbol{\rho}_t^{cm})$$

$$\boldsymbol{x}_t^{cm}, \boldsymbol{\rho}_t^{cm} = \mathbb{E}\Psi(s_t, \boldsymbol{x}_{t-1}^{cm}, \boldsymbol{\rho}_{t-1}^{cm})$$

Note how these conditional mean paths smooth the high frequency movements in the dynamics of the state variables but retain the low frequency drifts. As before, different lines correspond

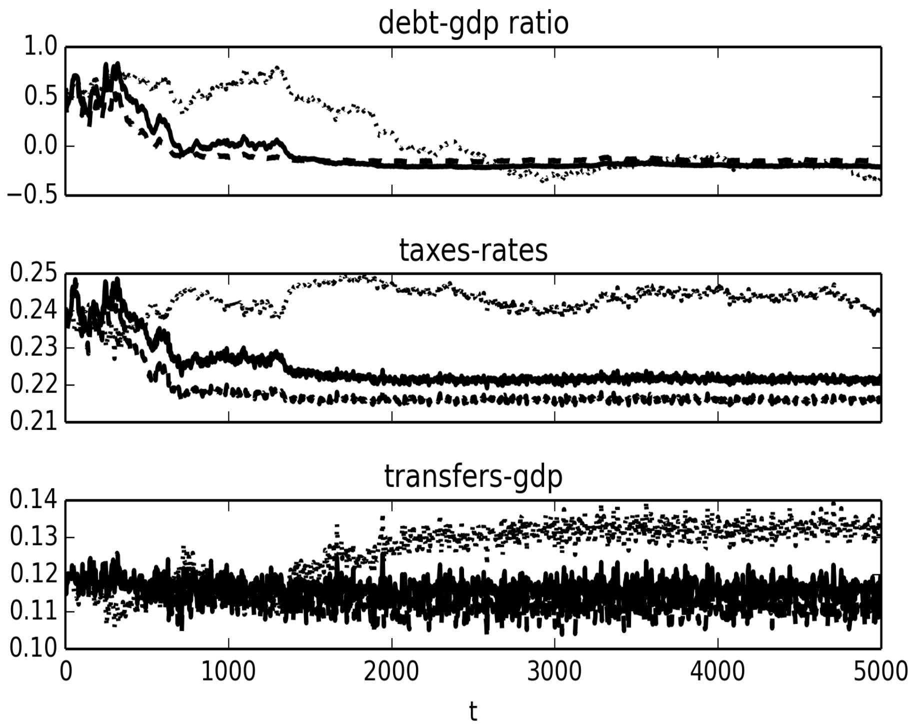


Figure 2: The dashed, bold and dotted lines plot simulations for a common sequence of shocks for different values of χ that generate $Corr[UcP, \epsilon] = \{-0.88, -0.60(\text{benchmark}), -0.22\}$ respectively

to different values of χ that general generate $Corr[UcP, \epsilon] \in \{-0.88, -0.60(\text{benchmark}), -0.22\}$. Thicker lines depict outcomes associated with larger values of χ in absolute magnitude. The figure shows that the speed of convergence is increasing and the magnitude of the limiting assets in decreasing in the strength of correlation between productivities and payoffs. This pattern confirms the approximation results characterized in Proposition 3.

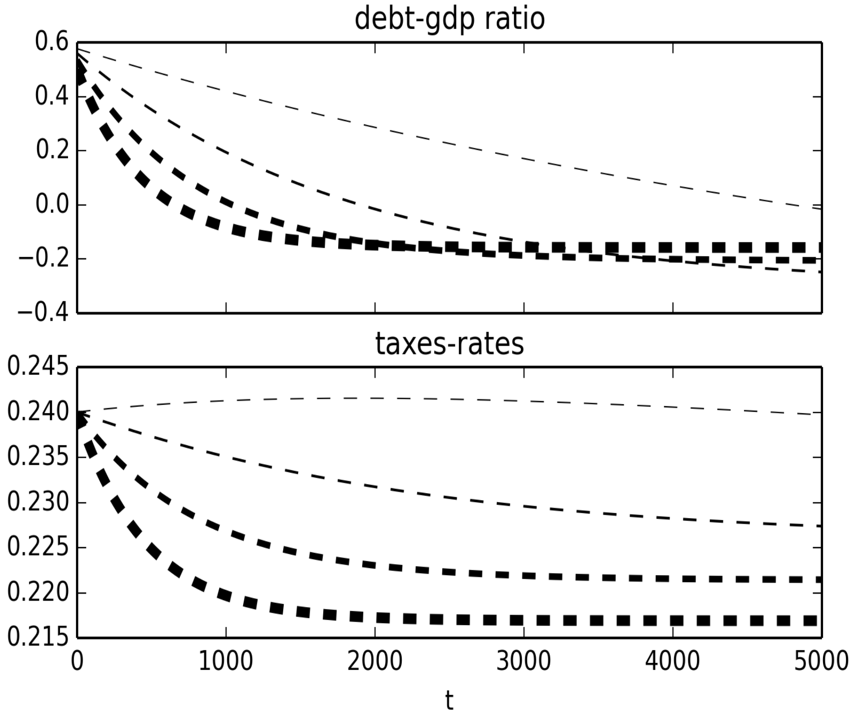


Figure 3: Conditional mean paths for different values of χ that generate $Corr[UcP, \epsilon] = \{-0.88, -0.79, -0.60(\text{benchmark}), -0.22\}$. Thicker lines representing more negative values.

Next we turn to the implications of Theorem 2 that the support of ergodic distribution can be very wide. To verify this we take the initial conditions at the end of the long simulation and subject the economy to a sequence of 100 periods of $\epsilon_{\theta,t}$ shocks that are 1 standard deviations below the mean. In figure 4 we see that given a sufficiently long sequence of negative productivity shocks the economy will eventually deviate significantly from its ergodic mean. However the tax rates spread very slowly (approximately 4 basis points per year) for the first 60 periods. From this we expect that the the ergodic distribution of taxes will be concentrated towards the mean. Theorem 3 and 4 suggest that the spread in tax rates depend on the size of the underlying shocks and how *far away* the economy from complete spanning. At our benchmark calibration for both reasons; redistributive concerns that imply the use of transfers to hedge aggregate shocks and fluctuations in endogenous returns that aid in spanning the required needs for revenues are important for the outcome that taxes do not spread too much.

Theorem 4 predicted that government assets B (in the long run) are decreasing in the redistributive motive of the government. We check this numerically here by plotting the long run assets of the government for different Pareto weights, indexed with a single parameter. Let

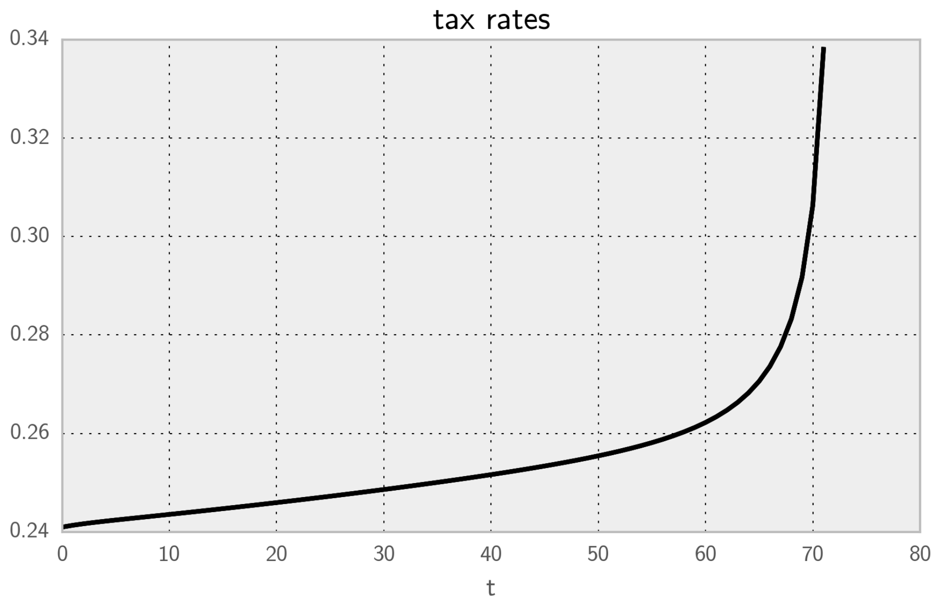


Figure 4: Taxes for a sequence of -1 s.d shocks to aggregate productivity of length 100

$\omega_1(\alpha) = \omega_1 + \alpha$ and $\omega_N(\alpha) = \omega_N - \alpha$. Increasing the parameter α shifts Pareto weights from the least productive agent towards the most productive agent relative to the benchmark $\alpha = 0$. In figure 5.2 we plot mean of the government assets in the ergodic distribution as a function of α and validate that higher values of α are associated with larger values long run government assets.

5.3 Short run

The analysis of the previous subsection studied very low frequency components of a Ramsey plan. Here we focus on business cycle frequencies to a) emphasize the role of inequality shocks and b) compare outcomes to actual US tax and debt policy. We compute the competitive equilibrium with tax-debt policy fitted to US, the optimal allocation with and without inequality, i.e., set the parameter $m = 0$ for two types of shock sequences:

1. A sequence that induces a one time one standard deviation impulse to $\epsilon_{\theta,t}$ and $\epsilon_{g,t}$ respectively.

2. A sequence $\{\epsilon_{\theta,t}, \epsilon_{g,t}\}$ recovered using data on aggregate labor earnings, tax rates, government expenditures and total debt for the period 1978-2010.

Figures 5.3 shows the impulse responses. The left panel plots responses to a one standard deviation fall in aggregate labor productivity $\epsilon_{\theta,t}$ in period 4 followed by no further productivity or expenditure shocks.

Under the benchmark calibration, the optimal policy (bold line) increases transfers and tax rates to counteract the effects of higher inequality in recessions. As compared to this the competitive equilibrium with US tax-debt policies(dashed line) implies a larger response in transfers but a decrease in tax rates. The optimal policy when we keep the relative skill distribution

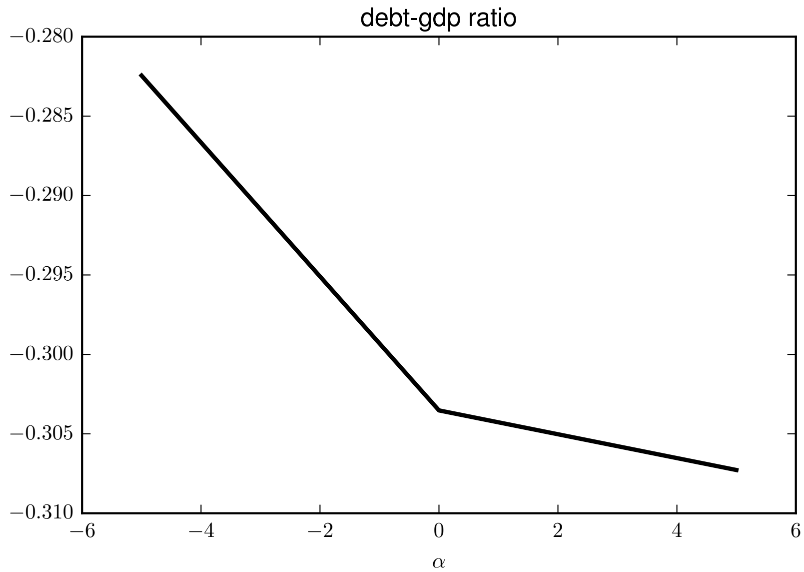


Figure 5: Long run debt-gdp ratio as a function of α: Higher α represent higher weights on productive agents.

unchanged (dotted line) has both taxes and transfers lower in recessions and the magnitude of change is smaller than the what is predicted under the benchmark calibration or the competitive equilibrium fitted to US tax-debt policies. The right panel of Figure 5.3 plots responses to a one standard deviation increase in government expenditure. Since relative skills are unchanged the tax and transfers policies under the optimal policy with and without inequality shocks are similar. We conclude that both the direction and magnitude of how taxes and transfers react to shocks is shaped by potential movements in inequality.

Next we filter the actual sequence of actual shocks $\{\epsilon_{\theta,t}, \epsilon_{g,t}\}$ using US data. To do this we use the policy rules from the solution to the competitive equilibrium with tax -debt policies fitted to U.S.. For all $t > 0$, given state variables $(B_{t-1}, \epsilon_{t-1})$ and the distribution of assets, we have a map from the realized shocks $(\epsilon_{B,t}, \epsilon_{\tau,t}, \epsilon_{g,t}, \epsilon_{\theta,t})$ to endogenous outcomes. Using HP filtered data on aggregate labor earnings, tax rates, government expenditures and total debt we invert this map to extract the values for these shocks . The flexibility of adding shocks to taxes, debt, and expenditure allows us extract the shocks that reconstruct all components of the government budget constraint and aggregate labor earnings in the competitive equilibrium. For simulating the Ramsey outcomes we only use the subset of shocks $\{\epsilon_{\theta,t}, \epsilon_{g,t}\}$ and compute the optimal taxes, transfers and debt. Figure 5.3 plots the recovered time series for $\{\epsilon_{g,t}, \epsilon_{\theta,t}\}$. Figures 5.3 compares the movements in debt, taxes and transfers for the optimal policy under the benchmark calibration with US data (left) and the optimal outcomes if we ignore inequality, i.e., set $m = 0$.

We find that the movements in debt under the optimal policy are in line with the data. The main differences are apparent in transfers and tax rates which are much less volatile under the optimal policy relative to what we see in data. Table 5.3 averages over sample paths of length 100 periods and reports the standard deviation, auto correlation, and correlation with exogenous

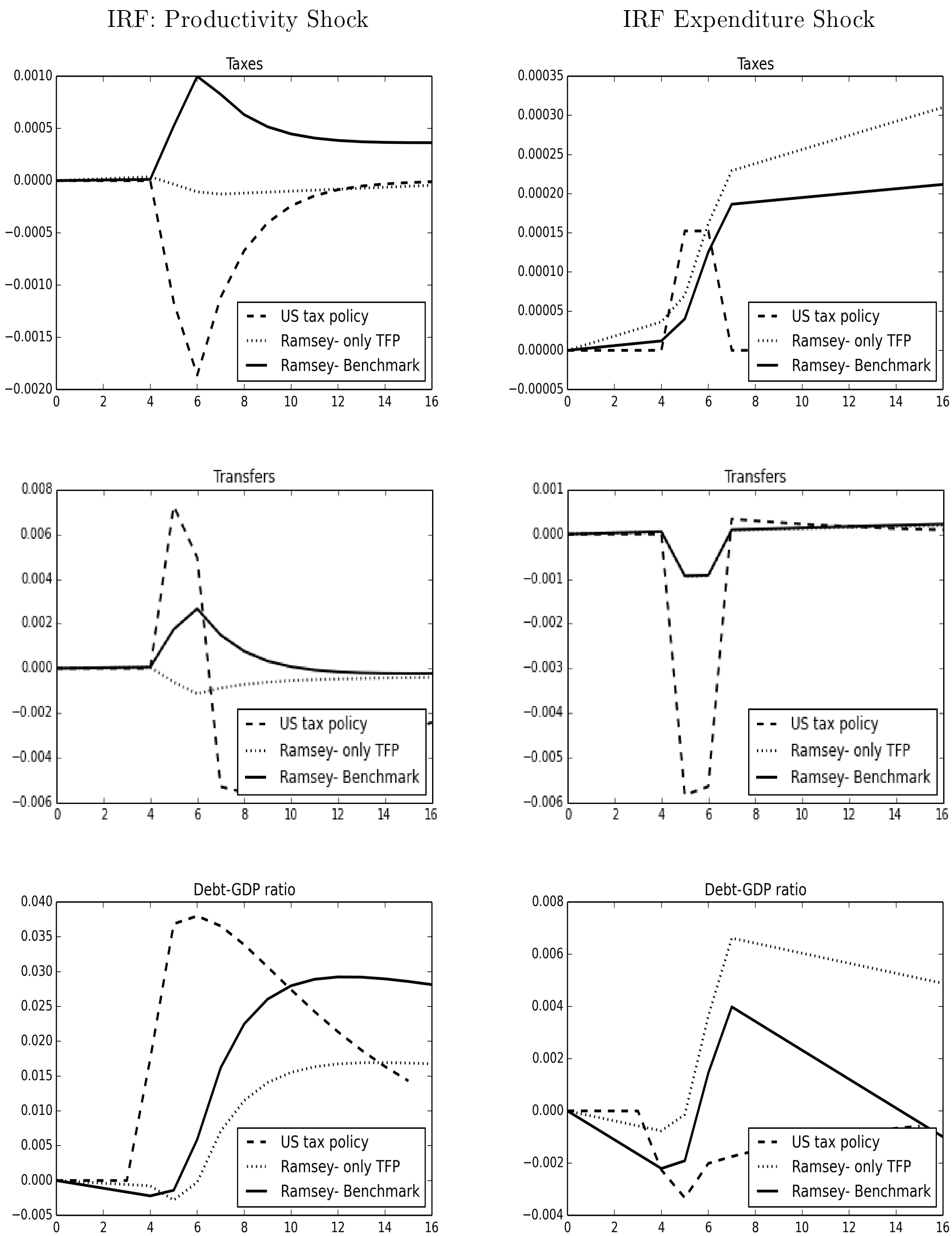


Figure 6: Impulse responses to productivity (left) and expenditure shocks (right).

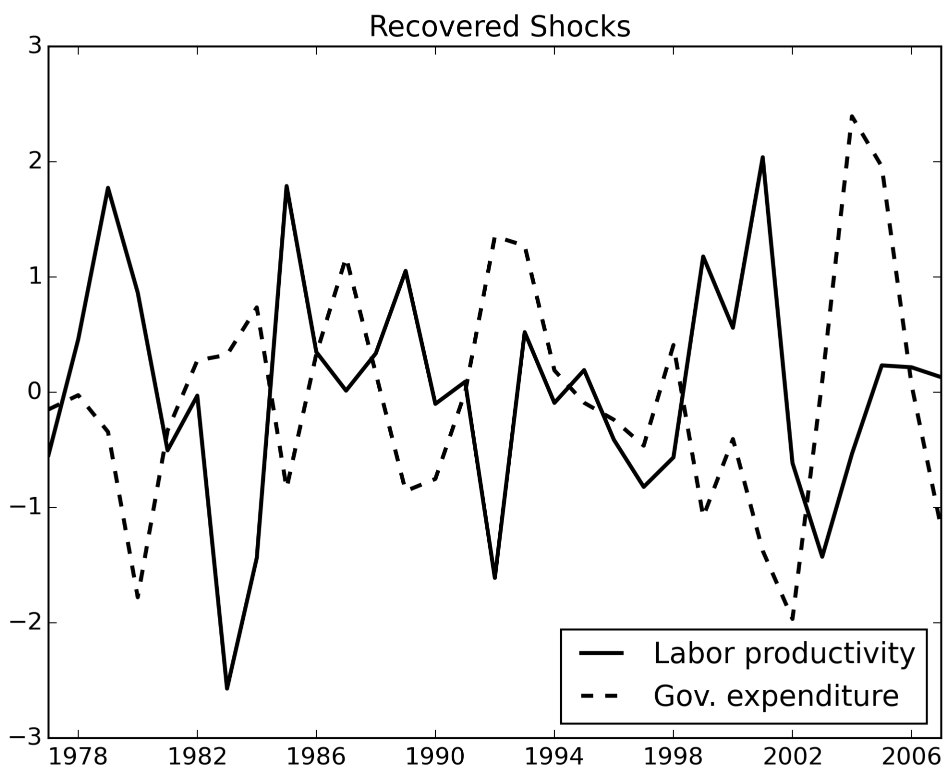


Figure 7: Recovered productivity and expenditure shocks

shocks for the tax rate and transfers. We see that under the optimal policy the movements in taxes and transfers are not only smaller but they have a much larger auto correlation. The optimal policy thus finances transitory fluctuations by small but long lasting changes in taxes and transfers.

Model	Competitive Equilibrium	Ramsey Benchmark	Ramsey No inequality	Data
Average Taxes	0.2399	0.2398	0.2400	0.2431
Average Transfers	0.1164	0.1183	0.1177	0.1069
S.D. Taxes	0.0062	0.0008	0.0004	0.0063
S.D. Transfers	0.0181	0.0008	0.0025	0.0109
Autocorrelation Taxes	0.0310	0.7475	0.8280	0.1893
Autocorrelation Transfers	0.0184	0.6677	0.6405	0.1814
Correlation Taxes and Output	0.2165	-0.7906	0.3218	0.1965
Correlation Transfers and Output	-0.1757	-0.4472	0.8505	-0.4046

Table 3: Sample moments for taxes and transfers averaged across simulations of 100 periods

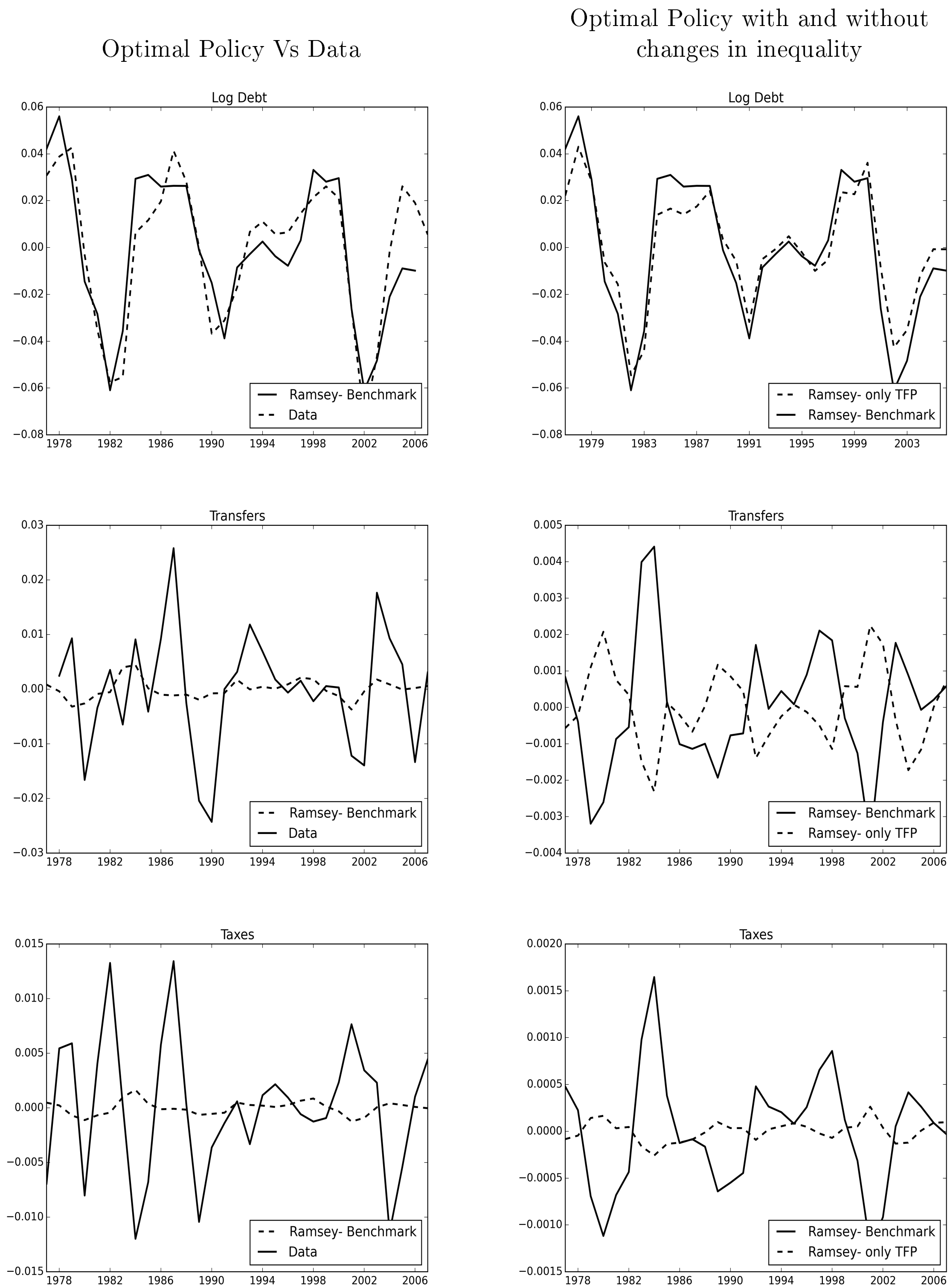


Figure 8: Outcomes along extracted shocks. The left panel compares the optimal policy under the benchmark calibration with data. The right panel compares the optimal policy with and without movements in inequality.

6 Concluding remarks

In a complete markets model like those of Lucas and Stokey (1983) and Chari et al. (1994), government debt is not a state variable. We have chosen to study a class of incomplete markets economies because the second and third questions posed in our first paragraph of this paper are interesting only in settings in which government debt is a state variable. The same fiscal-hedging motive that in complete markets economies eradicates government debt as a state variable also operates in our economy, but instead of wiping the slate clean each period, it now shapes a history-dependent plan for government debt. Another way that we depart from both the complete markets economies of Lucas and Stokey (1983) and Chari et al. (1994) and the incomplete markets economies of Aiyagari et al. (2002) and Farhi (2010) is that (a) we allow tax collections be affine rather than linear history-dependent functions of labor income, and (b) our economies are inhabited by agents heterogeneous in their human and financial assets. These two features make income redistribution through transfers join imperfect fiscal hedging as a force that contributes to the optimal dynamics of both government debt and the tax rate on labor income. They also add a smoothing motive along another dimension to the wish list of a Ramsey planner: in addition to an interest in smoothing fluctuations in the tax rate on labor income identified by Barro (1979), our Ramsey planner wants to smooth fluctuations in transfers, at least he does when he cares enough about poorer agents.

In the process of answering the three questions posed in our first paragraph, we have recast them by refining an appropriate notion of the state variable for an incomplete markets, heterogeneous agent economy like ours. The pertinent state variable is really the distribution of net assets across agents, not total government debt alone. Who owns government debt is as important as the total amount of government debt. We used this insight to rethink how the tax rate and transfers should respond to shocks at business cycle frequencies.

7 Appendix

7.1 Proofs of Proposition 1 and Corollary 1

Proof. Let $W\left(\{b_{i,-1}\}_i, B_{-1}\right)$ be the welfare in the best equilibrium associated with initial assets $(\{b_{i,-1}\}_i, B_{-1})$ and consider any two initial asset distributions $\left(\left\{b'_{i,-1}\right\}_i, B'_{-1}\right)$ and $\left(B''_{-1}, \left\{b''_{i,-1}\right\}_i\right)$ that satisfy (9). Suppose $W\left(\left\{b'_{i,-1}\right\}_i, B'_{-1}\right) > W\left(\left\{b''_{i,-1}\right\}_i, B''_{-1}\right)$. Let $\{\tau'_t, T'_t\}_t$ be the optimal policy and $\left\{\{c'_{i,t}, l'_{i,t}, b'_{i,t}\}_i, B'_t, R'_t\right\}_t$ be the optimal competitive allocation that attains $W\left(\left\{b'_{i,-1}\right\}_i, B'_{-1}\right)$. Consider the following policy $\hat{T}_0 = T'_0 + R'_0\left(b'_{1,-1} - b''_{1,-1}\right)$ and $\hat{T}_t = T'_t\ \forall t \geq 1$ and $\hat{\tau}_t = \tau'_t$ for all $t \geq 0$. We show that $\left\{\{c'_{i,t}, l'_{i,t}, b'_{i,t}\}_i, B'_t, R'_t\right\}_t$ is a competitive equilibrium for $\left(\left\{b''_{i,-1}\right\}_i, B''_{-1}\right)$ given $\{\hat{\tau}_t, \hat{T}_t\}_t$, which implies $W\left(\left\{b'_{i,-1}\right\}_i, B'_{-1}\right) \leq W\left(\left\{b''_{i,-1}\right\}_i, B''_{-1}\right)$.

The allocation $\left\{\{c'_{i,t}, l'_{i,t}, b'_{i,t}\}_i, B'_t, R'_t\right\}$ satisfies all the feasibility constraints by construction, so it remains to show that $\{c'_{i,t}, l'_{i,t}, b'_{i,t}\}_i$ remains the optimal choice for each consumer i. Suppose that some other $\left\{\hat{c}_{i,t}, \hat{l}_{i,t}, \hat{b}_{i,t}\right\}_t$ gives agent i strictly higher utility than $\{c'_{i,t}, l'_{i,t}, b'_{i,t}\}_i$. Then the choice $\left\{\hat{c}_{i,t}, \hat{l}_{i,t}, \hat{b}_{i,t}\right\}_t$ is also feasible for consumer i who has initial assets $b'_{i,-1}$ and faces taxes $\{\tau'_t, T'_t\}_t$, which contradicts the assumption that $\{c'_{i,t}, l'_{i,t}, b'_{i,t}\}_i$ is the optimal choice in the economy with taxes $\{\tau'_t, T'_t\}_t$ and assets $\left(\left\{b'_{i,-1}\right\}_i, B'_{-1}\right)$. This proves Proposition 1.

To prove Corollary 1, let,

$$\hat{T}_t = T_t + \left(\hat{b}_{1,t} - b_{1,t}\right) - R_t\left(\hat{b}_{1,t-1} - b_{1,t-1}\right) \text{ for all } t \geq 0.$$

Since $\left\{\hat{b}_{i,t}\right\}_{i,t}$ satisfies 10 and is bounded,

$$\hat{b}_{i,t} = \mathbb{E}_t \sum_{T\geq t}\left(\frac{c(s^T) - (1-\tau(s^T))\theta_i(s^T)l_i(s^T) - \hat{T}(s^T)}{\prod_{k=0}^{k=T-t} R_k}\right) \geq \inf_{s^\infty} \sum_{T\geq t, s^T \in s^\infty} \frac{-(1-\tau(s^T))\theta_i(s^T)l_i(s^T) - \hat{T}(s^T)}{\prod_{k=0}^{k=T-t} R_k}$$

and $\left\{\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\}_i, \hat{B}_t, R_t, \tau_t, \hat{T}_t\right\}_t$ satisfies the natural debt limits. Then the same steps as above imply that $\left\{\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\}_i, \hat{B}_t, R_t\right\}_t$ is a competitive equilibrium for $\left\{\hat{\tau}_t, \hat{T}_t\right\}_t$ and $\left(\{b_{i,-1}\}_i, B_{-1}\right)$. □

7.2 Proof of Lemma 1

Proof. Let $\mu(s), \phi(s), \lambda(s)$ be the Lagrange multipliers on the constraints (14a), (14b), (14c). We take the first order conditions of with respect to $c_1(s), c_2(s), l_1(s)$.

$$\omega - \mu(s) = n\phi(s), \tag{30a}$$

$$1 - \omega + \mu(s) - \phi(s)(1-n) + \lambda(s) = 0, \tag{30b}$$

$$-\omega l^\gamma(s) + \mu(s)(1+\gamma)l^\gamma(s) + n\phi(s)\theta = 0. \tag{30c}$$

To prove Lemma 1, we show that $\frac{\omega}{n} > \frac{1+\gamma}{\gamma}$ is sufficient for the Lagrange multiplier $\lambda(s) \geq 0$ on the non-negativity constraint (14c) to be strictly positive. Summing (30a) and (30b) we establish that $\phi(s) \geq 1$. Therefore (30c) implies that $\mu(s) \leq \frac{\omega}{1+\gamma}$. Use these bounds in equation (30b) to show that if $\omega - n - \frac{\omega}{1+\gamma} > 0$ then $\lambda(s) > 0$ for all s. The last inequality can be re-written as $\omega > n\left(\frac{1+\gamma}{\gamma}\right)$. If it holds, then $c_2(s) = 0$ for all s. Agent 2 budget constraint then implies that $T_t = 0\ \forall t \geq 0$. □

7.3 Proof of Proposition 2

Proof. For $\omega > \bar{\omega}$ we can simplify the problem (13). We first show that the value function $V(B\_)$ is strictly concave and use this to derive some properties of policy rules $B(s, B\_)$ for arbitrary payoffs P to prove Part 1 and Part 2.

Lemma 1. *$V(B\_)$ is strictly concave and differentiable for interior $B\_$.*

Proof. Transform the variable l such that $L \equiv l^{1+\gamma}$. We will recast problem[14] where the maximization problem using L and show that the problem is convex. Substituting for $nc_1(s)$ we get,

$$V(B\_) = \max_{L(s),B(s)} \sum_{s\in S} \pi(s) \left[\frac{n\gamma}{1+\gamma} L(s) - \frac{1}{\beta} P(s) B\_ + B(s) + \beta V(B(s)) \right], \tag{31}$$

subject to

$$B(s) - \frac{1}{\beta} P(s) B\_ + g(s) \leq n\theta L(s)^{\frac{1}{1+\gamma}} - nL(s), \tag{32a}$$

$$\underline{B} \leq B(s) \leq \bar{B}.$$

Let $\phi(s)$ be the multiplier constraint (32a). The first order condition with respect to $L(s)$ is

$$\frac{n\gamma}{1+\gamma} + n\phi(s) \left[\frac{\theta}{1+\gamma} L(s)^{\frac{-\gamma}{1+\gamma}} - 1 \right] = 0. \tag{33}$$

If $\phi(s) = 0$, the optimal $L(s)$ solves $\max_{L(s)\geq 0} \pi(s) \frac{n\gamma}{1+\gamma} L(s)$ which sets $L(s) = \infty$. This violates feasibility for any finite $B\_, B(s)$. This implies $\phi(s) > 0$ and consequently,

$$L(s)^{\frac{\gamma}{1+\gamma}} \geq \frac{\theta}{1+\gamma}. \tag{34}$$

Inequality (34), which is a lower bound for $L(s)$, simply means that it is not optimal to set tax rates to the right of the peak of the Laffer curve. In this region, the right hand side of equation (32a) (which are the total tax revenues) is decreasing as a function of $L(s)$. We can therefore define a function $\Phi(.)$ such that

$$\Phi\left(\frac{n\gamma}{1+\gamma} L \right) = n\theta L^{\frac{1}{1+\gamma}} - nL. \tag{35}$$

We next verify that $\Phi(.)$ is decreasing and strictly concave in the interior. To see this note $\Phi'(.)\frac{\gamma}{1+\gamma} = \left[\frac{\theta}{1+\gamma} L(s)^{\frac{-\gamma}{1+\gamma}} - 1 \right] < 0$ and $\Phi''(.)n\left(\frac{\gamma}{1+\gamma}\right)^2 = \left[\left(\frac{\theta}{1+\gamma}\right) \left(\frac{-\gamma}{1+\gamma}\right) L(s)^{-\left(\frac{2\gamma+1}{1+\gamma}\right)} \right] < 0$. Let $D = \Phi\left(\frac{n\gamma}{1+\gamma} \left(\frac{\theta}{1+\gamma}\right)^{\frac{1+\gamma}{\gamma}} \right)$ be the maximum revenue the government can raise. We can rewrite the problem (31) as

$$V(B\_) = \max_{B(s)} \sum_{s\in S} \pi(s) \left[\Phi^{-1}\left(B(s) - \frac{1}{\beta} P(s) B\_ + g(s) \right) + B(s) - \frac{1}{\beta} P(s) B\_ + \beta V(B(s)) \right] \tag{36}$$

subject to

$$B(s) - \frac{1}{\beta} P(s) B\_ + g(s) \leq D, \tag{37a}$$

$$\underline{B} \leq B(s) \leq \bar{B}. \tag{37b}$$

[14] We allow the resource constraint to hold with a weak inequality. This allows us to show the desired properties of $V(B\_)$ and we will also show that in this general problem the inequality constraint always binds. Hence the solutions to both the problems, with or without the weak inequality are identical.

Since $\Phi(.)$ is concave and strictly decreasing, the period utility function in objective (36) is strictly concave too. Standard arguments as in Proposition 4.6-4.8 in Stokey, Lucas, and Prescott (1989) apply and we conclude that $V(B\_)$ is strictly concave in the interior. Differentiability at interior points comes from Benveniste-Scheinkman (1979). $\square$

We list the first order condition with respect to $B(s)$ for (36) and will refer to them for the rest of the proof:

$$\Phi^{-1\prime}\left(B(s) - \frac{1}{\beta}P(s)B\_ + g(s)\right) + 1 + \beta V'(B(s)) - \bar{\kappa}(s) + \underline{\kappa}(s). \tag{38}$$

The multiplier on the constraint (37a) is zero for all s. This follows from the definition of D as the tax revenues associated with the highest tax the government can raise. The implied $L(s)$ violates equation (33).

The next lemma uses the Lemma 1 to characterize some properties of the policy rules $B(s, B\_)$.

Lemma 2. *Let $B(s, B\_)$ be the optimal policy for assets. It has the following properties:*

- *$B(s, B\_)$ is strictly increasing in $B\_$ for all s where $B(s, B\_)$ is interior.*
- *For any interior $B\_ \in (\underline{B}, \bar{B})$, there are s, s' s.t. $B(s, B\_) \geq B\_ \geq B(s', B\_)$. Moreover, if there are any states s'', s''' s.t. $B(s'', B\_) \neq B(s''', B\_)$ for some $B\_$ those inequalities are strict.*

Proof. Suppose $B' < B''$ but $B\left(s, B'\right) \geq B\left(s, B''\right)$. Strict concavity of $V(.)$ implies,

$$V'\left(B\left(s, B''\right)\right) \geq V'\left(B\left(s, B'\right)\right).$$

At an interior solution, the first order condition with respect to $B(s)$ from the formulation in (36) implies that

$$\Phi^{-1\prime}\left(B\left(s, B''\right) - \frac{1}{\beta}P(s)B'' + g(s)\right) \leq \Phi^{-1\prime}\left(B\left(s, B'\right) - \frac{1}{\beta}P(s)B' + g(s)\right),$$

$$B\left(s, B''\right) - B\left(s, B'\right) \geq \frac{1}{\beta}P(s)[B'' - B'].$$

The last inequality yields a contradiction and this shows that $B(s, B\_)$ is increasing.

Equation (38) together with the envelope proposition gives us

$$\mathbb{E}P(s)\beta V'(B(s)) = \beta V'(B\_) + \mathbb{E}P(s)\bar{\kappa}(s) - \mathbb{E}P(s)\underline{\kappa}(s).$$

We can rewrite this as

$$\tilde{\mathbb{E}}\beta V'(B(s)) = \beta V'(B\_) - \underline{\kappa} + \bar{\kappa}, \tag{39}$$

with $\tilde{\pi}(s) = P(s)\pi(s)$, $\underline{\kappa} = \tilde{\mathbb{E}}\underline{\kappa}(s)$, and $\bar{\kappa} = \tilde{\mathbb{E}}\bar{\kappa}(s)$.

First we show that $\exists s$ such that $B(s, B\_) \geq B\_$. Suppose not, then $B(s, B\_) < B\_$ for all s. Strict concavity of V implies

$$\tilde{\mathbb{E}}V'(B(s, B\_)) > V'(B\_).$$

Since $B(s, B\_) < B\_ < \overline{B}$, the multipliers $\overline{\kappa} = 0$ and (39) implies

$$\tilde{\mathbb{E}}V'(B(s,B\_)) \leq V'(B\_).$$

This yields a contradiction.

Similar arguments show that if there is at least one $B\,(s,B\_)$ s.t. $B\,(s,B\_) > B\_$, then there must be some s' s.t. $B\,(s',B\_) < B\_$. For an arbitrary $B\_$, either we have $B(s,B\_) = B\_$ for all s or there exists a s such that $B(s,B\_) > B\_$. Under the condition that there exists s'', s''' such that $B(s'',B\_) \neq B(s''',B\_)$ the former cannot hold. This shows that the inequalities can be made strict. $\square$

Let $\mu(s,B\_) = -\beta V'(B(s,B\_))$, our next lemma orders $\mu(s,B\_)$ relative to $P(s)$.

Lemma 3. *For a pair of shocks s', s'' if $P(s') > P(s'')$ then there exists a $B^*(s's'')$ such that for all $B\_ > (<) B^*(s',s'')$ we have $\mu(s',B\_) > (<)\ \mu(s'',B\_)$. If $\underline{B} < B^*(s's'') < \overline{B}$ then $\mu(s',B^*(s's'')) = \mu(s'',B^*(s's''))$.*

Proof. Define $B^*(s',s'') = \frac{\beta[g(s')-g(s'')]}{P(s')-P(s'')}$. For an interior solution, the first order condition with respect to $B(s)$ requires

$$\Phi^{-1'}\left(B(s) - \frac{1}{\beta}P(s)B\_ + g(s)\right) + \beta V'(B(s)) + 1 = 0.$$

Concavity of Φ^{-1} and V implies that the left hand side of this equation is a decreasing function of $B(s)$. If $B\_ > (<) B^*(s',s'')$ then $g(s') - \frac{1}{\beta}P(s')B\_ < (>) g(s'') - \frac{1}{\beta}P(s'')B\_$, which immediately implies that $B(s',B\_) > (<) B(s'',B\_)$. As V is strictly concave we conclude that $\mu(s',B\_) > (<)\ \mu(s'',B\_)$. Finally if $\underline{B} < B^*(s',s'') < \overline{B}$ then continuity and monotonicity of the function $\mu(s',B\_)$ with respect to $B\_$, $\mu(s',B\_) - \mu(s'',B\_)$ is zero uniquely at $B\_ = B^*(s's'')$. $\square$

Now we will use these lemmas to prove the proposition. Lemma 2 states that $B(s,B\_)$ is increasing. Given that shocks are i.i.d, the existence of a invariant distribution of B_t follows from Hopenhayn and Prescott (1992) (see corollary 5).

To show Part 1, note that $P \in \mathcal{P}^*$ is necessary and sufficient for existence of B^* such that $B(s,B^*) = B^*$ for all s. The necessary part follows from taking differences of the (32a) for s',s'', which yields

$$[P(s'') - P(s')]\frac{B^*}{\beta} = g(s') - g(s'').$$

Thus $P \in \mathcal{P}^*$. The sufficient part follows from the Lemma 3. If $P \notin \mathcal{P}^*$, then $B^*(s',s'')$ that defines the pairwise crossing will not be same across all pairs s', s''. Thus B^* that satisfies $B(s,B^*)$ independent of s will not exist.

Let $\mu\_ = -V'(B\_)$, for . Equation (39) implies,

$$\mu\_ = \mathbb{E}P(s)\mu(s) + \underline{\kappa}(s). \tag{40}$$

By decomposing $\mathbb{E}\mu(s)P(s)$ in equation (40), we obtain (using $\mathbb{E}P(s) = 1$)

$$\mu\_ = \mathbb{E}\mu(s) + \mathrm{cov}(\mu(s), P(s)) + \underline{\kappa}(s). \tag{41}$$

Lemma 3 implies that for $B\_ > B^*$, the covariance term $\mathrm{cov}(\mu(s), P(s)) > 0$. Thus

$$\mu_t \geq \mathbb{E}_t\mu_{t+1}.$$

In this region μ_t is a super martingale. Denote $\mu^* \equiv -V'(B^*)$ and note that,

$$\frac{\partial \mu(s,B\_)}{\partial \mu(B\_)} = \underbrace{\frac{\partial \mu(s,B\_)}{\partial B(s,B\_)}}_{>0} \underbrace{\frac{\partial B(s,B\_)}{\partial B\_}}_{>0} \underbrace{\left(\frac{\partial \mu(B\_)}{\partial B\_}\right)^{-1}}_{>0} > 0.$$

This implies $\mu_t > \mu^* \implies \mu_{t+1} > \mu^*$. Thus the super martingale is bounded below by μ^*. Using standard martingale convergence proposition we can conclude that μ_t converges . The uniqueness of steady state implies that it can only converge to μ^*. Since $\mu(s,B\_)$ is continuous and monotonic in $B\_$, as μ_t converges to μ^*, B_t converges to B^*. The argument for the $B_{-1} < B^*$ is symmetric and we omit it for brevity. From equation (38), as the μ_t converges, labor supply converges too and the tax rate is eventually constant.
For Part 2, when $P^* \notin \mathcal{P}^*$, there is no interior absorbing $B\_$ such that $B(s,B\_) = B\_$. Using Lemma 2, from any $B\_$, we can construct a finite sequence of shocks such that $B(s^T|B\_) > \bar{B} - \epsilon$ and some other finite sequence of shocks such that $B(\tilde{s}^T|B\_) < \underline{B} + \epsilon$. Thus there is a T' that is equal to the maximum length of s^T and $\tilde{s}^T$ across $B\_$ such that both the thresholds are crossed in finite steps. For $\epsilon > 0$ and B_{-1}, let $A_T = \{B_t < \underline{B} + \epsilon \text{ and } B_{t'} > \bar{B} - \epsilon \text{ for some } t, t' \leq T\}$ and let $k(n) = nT'$. Since $A_T \subset A_{T+1}$, a sufficient condition for $\Pr\{A_T \text{ occurs i.o }\} = 1$ is that $\sum_n \Pr\{A_{k(n)+1}|A_{k(n)}\}$ diverges. From the previous discussion we can see that $\Pr\{A_{k(n)+1}|A_{k(n)}\}$ is (uniformly) bounded from below and hence the $\sum_n \Pr\{A_{k(n)+1}|A_{k(n)}\} = \infty$.
For the last implication of Part 2, let B_1, B_2 be defined by

$$B_1 = \min_{s',s''} \left\{B^*(s',s'')\right\},$$

$$B_2 = \max_{s',s''} \left\{B^*(s',s'')\right\}.$$

For $B\_ > B_2$ Lemma 3 showed that $P(s) > P(s')$ implies $\mu(s,B\_) > \mu(s',B\_)$ and hence $\text{cov}(\mu(s), P(s)) > 0$. From equation (41), we have

$$\mu\_ > \mathbb{E}\mu(s,B\_).$$

If P is sufficiently volatile, $P(s'') - P(s') > \frac{\beta(g(s'')-g(s'))}{\underline{B}}$ and therefore

$$B_1 = \min\left\{B^*_{s',s''}\right\} > \underline{B}$$

Similar arguments imply that for $B\_ < B_1$ the $\text{cov}(\mu(s), P(s)) < 0$ and

$$\mu\_ < \mathbb{E}\mu(s,B\_).$$

For the implications on the tax rates, note from equation (38) we have,

$$-\mu(s,B\_) = -\Phi^{-1'}\left[\theta n L(s,B\_)^{\frac{1}{1+\gamma}} - nL(s,B\_)\right] - 1. \tag{42}$$

This defines $L(s,B\_) = \mathcal{L}(\mu(s,B\_))$ and the implicit function proposition implies,

$$-1 = -n\Phi^{-1''}(.)\left[\frac{\theta}{1+\gamma}\mathcal{L}^{\frac{-\gamma}{1+\gamma}} - 1\right]\mathcal{L}'.$$

Since $\Phi^{-1}(.)$ is concave, and $\left[\frac{\theta}{1+\gamma}L(s)^{\frac{-\gamma}{1+\gamma}}-1\right]<0$ for the relevant range for L, we can conclude that $\mathcal{L}'>0$. The household's optimal choice of labor implies $\tau(s,B\_)=1-\frac{L(s,B\_)^{\frac{\gamma}{1+\gamma}}}{\theta}$. We can define a function $\hat{\tau}(B)$ by

$$\hat{\tau}(B)=1-\theta^{-1}\mathcal{L}[-V'(B)]^{\frac{\gamma}{1+\gamma}}, \tag{43}$$

and $\hat{\tau}'=\frac{\theta\gamma}{1+\gamma}\mathcal{L}^{-\frac{1}{1+\gamma}}\mathcal{L}'V''<0$.

□

7.4 Proof of Proposition 3

To facilitate our approximations required to prove Proposition 3, we will use a recursive representation of the optimal allocation using state variable $\mu\_$ where $\mu\_=-\beta V'(B\_)$. Given some payoff vector P, we solve for $B(\mu\_,P)$ and $\mu(s,\mu\_,P)$ that satisfy the following equations for all $\mu\_$,

$$B(\mu(s,\mu\_,P),P)=\Phi\left(\frac{n\gamma}{1+\gamma}\mathcal{L}[\mu(s,\mu\_,P)]\right)-g(s)+\frac{P(s)}{\beta}B(\mu\_,P), \tag{44}$$

$$\mu\_=\mathbb{E}\mu(s,\mu\_,P)P(s). \tag{45}$$

We will explicitly recognize the fact that the unknown functions $B(.)$ and $\mu(.)$ depend on the payoff vector P. The functions $\Phi(.)$ and $\mathcal{L}(.)$ are defined in equations (35) and (42) in the proof of Proposition 2. With a slight abuse of notation, we will use $\Phi(\mu)$ to mean $\Phi\left(\frac{\gamma}{1+\gamma}\mathcal{L}[\mu]\right)$.

We will approximate these functions around around economy with payoffs in $\bar{P}\in\mathcal{P}^*$. For $\bar{P}\in\mathcal{P}^*$ and an associated $\bar{B}(\bar{P})$ we can define $\bar{\mu}(\bar{P})=-V'(\bar{B}(\bar{P}))$.[15] First we differentiate (44) and (45) to solve for the derivatives of the the functions $\mu(s,\mu\_,P)$ and $B(\mu\_,P)$ with respect to $(\mu\_,P)$ at $(\bar{\mu}\left(\bar{P}\right),\bar{P})$ where $\bar{P}\in\mathcal{P}^*$.[16] Next we will use these expressions to compute the mean and variance of the ergodic distribution associated with the approximated policy rules and finally as a last step we propose a particular choice of the point of approximation to get the expressions in Proposition 3.

From now, we drop P from the arguments of the functions to keep the notation simple. Differentiating equations (44) and (45) with respect to $\mu\_$ around $(\overline{\mu},\overline{P})$ we obtain

$$\frac{\partial B(\bar{\mu})}{\partial\mu\_}=\frac{\Phi'(\overline{\mu})}{\frac{\beta}{\bar{B}^2}\sigma_g^2+\frac{1-\beta}{\beta}},$$

$$\frac{\partial\mu'(s,\bar{\mu})}{\partial\mu\_}=\frac{\overline{P}(s)}{\frac{\beta^2}{\bar{B}^2}\sigma_g^2+1}=\frac{\overline{P}(s)}{\mathbb{E}\overline{P}(s)^2}.$$

We can perform the same steps for and differentiate with respect $P(s)$ for all s.

$$\frac{\partial B(\bar{\mu})}{\partial P(s)}=-\pi(s)\bar{B}\left(\frac{\mathbb{E}\overline{P}(s)^2-\bar{P}(s)}{\mathbb{E}\overline{P}(s)^2}\right),$$

[15] $\bar{\mu}(\bar{P})$ also solves the equation $\bar{B}(\bar{P})=\frac{\beta}{1-\beta}\left(\mathbb{E}g-\Phi(\bar{\mu})\right)$.

[16] Usually perturbation approaches to solve equilibrium conditions as above look for the solutions to $\{\mu(s|\mu\_)\}$ and $b(\mu\_)$ around deterministic steady state of the model. However the first order expansion of $\mu(s,\mu\_)$ will imply that it is a martingale. Such approximations are not informative about the ergodic distribution.

$$\frac{\partial \mu(s',\bar{\mu})}{\partial P(s)} = \frac{-\bar{B}}{\beta\left[\Phi'(\bar{\mu}) - \frac{\partial B}{\partial \mu}\right]}\left(1_{s,s'} - \frac{\pi(s)\bar{P}(s)\bar{P}(s')}{\mathbb{E}\bar{P}^2}\right).$$

Using the derivatives that we computed at $\left(\bar{\mu},\bar{P}\right)$, we can characterize the dynamics of $\hat{\mu}_t \equiv \mu_t - \overline{\mu}$ using the approximation:

$$\hat{\mu}_{t+1} = \frac{\partial \mu(s_{t+1},\bar{\mu})}{\partial \mu_-}\hat{\mu}_t + \frac{\partial \mu(s_{t+1},\bar{\mu})}{\partial P(s_{t+1})}[P(s_{t+1}) - \bar{P}(s_{t+1})]. \tag{48}$$

Let $D(s)$ and $C(s)$ be the respective coefficients and we express equation (48) as:

$$\hat{\mu}_{t+1} = D(s_{t+1})\hat{\mu}_t + C(s_{t+1}). \tag{49}$$

Note that both $D(s)$ and $C(s)$ are random variables and denote their means $\overline{D}$ and $\overline{C}$, and variances σ_D^2 and σ_C^2 . Taking expectations of equation (49), we solve for the ergodic mean $\mathbb{E}\hat{\mu}$,

$$\mathbb{E}\hat{\mu} = \frac{\overline{C}}{1-\overline{D}},$$

and analogously for the ergodic variance $\sigma_{\hat{\mu}}^2$ we get

$$\sigma_{\hat{\mu}}^2 = \frac{\sigma_D^2(\mathbb{E}\hat{\mu})^2 + \sigma_{DC}\mathbb{E}\hat{\mu} + \sigma_C^2}{1-\overline{D}^2 - \sigma_D^2}.$$

Using the ergodic means and variances of $\hat{\mu}$, we can obtain those for assets $\hat{B}_t \equiv B_t - \bar{B}$ with $\mathbb{E}\hat{B} = \frac{\partial B(\bar{\mu})}{\partial \mu_-}\mathbb{E}\hat{\mu}$ and $\sigma_{\hat{B}}^2 = \sigma_{\hat{\mu}}^2\left(\frac{\partial B(\bar{\mu})}{\partial \mu_-}\right)^2$. To get the expressions in Proposition 3, we solve the least square projection (17). Since payoffs in $\bar{P} \in \mathcal{P}^*$ are associated with a unique $\bar{B}(\bar{P})$ and $\bar{\mu}(\bar{P})$, we can define a mapping $\mathbf{P}[s,\bar{\mu}(\bar{P})] = \bar{P}(s)$ and rewrite the problem as choosing $\tilde{\mu}$ so as to minimize the variance of the difference between P and the set $\mathcal{P}^*$. Let P^* be the solution to this minimization problem. The first order condition gives us

$$2\sum_{s'}\pi(P(s') - P^*(s'))\frac{\partial \mathbf{P}(s,\mu^*)}{\partial \mu} = 0. \tag{50}$$

Since $P^*(s) = 1 + \frac{\beta}{B^*(\mathbf{P}[s,\mu^*])}(g(s) - \mathbb{E}g)$, we have $\frac{\partial \mathbf{P}(s,\mu^*)}{\partial \mu} \propto P^*(s) - 1$. Substituting back in equation (50), we see that the optimal choice of μ^* is such that

$$\mathbb{E}\left[(P-P^*)P^*\right] = 0. \tag{51}$$

At the optimal P^*, coefficients $\{C(s)\}_s$ of the linearized system (49) are given by

$$C(s) = -\frac{B^*}{\beta\left[\Phi'(\mu^*) - \frac{\partial B(\mu^*)}{\partial \mu_-}\right]}\left[P(s) - P^*(s)\right].$$

Taking expectations, we get that $\bar{C} = 0$. Thus the linearized system will have the same ergodic mean for μ_t as $\mu^*(P^*)$ and $\mathbb{E}B = B^*(P^*)$. The expressions for ergodic variance of government assets simplifies to

$$\sigma_B^2 = \frac{(B^*)^2}{\mathbb{E}\left(P^*\right)^2\mathrm{var}(P^*)}\|P-P^*\|^2.$$

Noting that $\mathbb{E}\left(P^*\right)^2 = 1 + \text{var}(P^*) > 1$, we obtain the bound

$$\frac{\sigma_B}{B^*} \leq \sqrt{\frac{\|P - P^*\|^2}{\text{var}(P^*)}}.$$

Lastly we derive the rates at which the conditional mean converges to its ergodic values. Taking expectations of (49) we get that $\frac{\mathbb{E}\hat{\mu}_{t+1}}{\hat{\mu}_t} = \overline{D}$, where we have used the result that $\overline{C} = 0$. We now express $\overline{D}$ in terms of the primitives P, g :

$$\overline{D} = \frac{1}{\mathbb{E}\left(P^*\right)^2}.$$

Using the orthogonal decomposition from equation (51) and the expression $P^*(s) - 1 = \frac{\beta}{B^*}[g(s) - \mathbb{E}g]$ we get,

$$\overline{D} = \frac{1}{1 + \text{var}(P)\text{corr}^2(P, g)}.$$

Since $\hat{B}_t$ is linear in $\hat{\mu}_t$, the ergodic mean of assets (under the approximated law of motion) also converges at the same rate.

7.5 Proof of Proposition 4

Proof. Using the transformation $L \equiv l_1^{1+\gamma}$ and function $\Phi(.)$ defined in (35) we can write the Bellman equation for $\omega \leq \bar{\omega}$ as

$$V\left(B\_\right) = \max_{B(s),T(s)} \mathbb{E}\left[\omega\Phi^{-1}\left(B(s) - \frac{1}{\beta}P\left(s\right)B\_ + g\left(s\right) + T(s)\right) + \omega\left(B(s) - \frac{1}{\beta}P\left(s\right)B\_\right) + nT(s) + \beta V\left(B\left(s\right)\right)\right] \tag{52}$$

subject to

$$B(s) - \frac{1}{\beta}P\left(s\right)B\_ + g\left(s\right) + T(s) \leq D, \tag{53a}$$

$$T(s) \geq 0, \tag{53b}$$

$$\underline{B} \leq B\left(s\right) \leq \bar{B}. \tag{53c}$$

The period gain function is concave in the arguments $\{B(s), T(s)\}$, and the choice set is compact. Using standard arguments and steps as in Lemma 1 we can show that V is concave. Let $\pi(s)\lambda_1(s), \pi(s)\lambda_2(s)$ be the Lagrange multipliers on constraints (53a) and (53b) and $\pi(s)\underline{\kappa}(s), \pi(s)\bar{\kappa}(s)$ on the bounds in constraints (53c). We list the first order conditions for problem (52) with respect to $B(s), T(s)$ and the envelope conditions and refer to them in the rest of the proof of the proposition:

$$\omega\Phi^{-1'}\left(B(s) - \frac{1}{\beta}P\left(s\right)B\_ + g\left(s\right) + T(s)\right) + n - \lambda_1(s) + \lambda_2(s) = 0, \tag{54a}$$

$$\omega - n - \lambda_2(s) - \bar{\kappa}(s) + \underline{\kappa}(s) = -\beta V'(B(s)), \tag{54b}$$

$$\beta V'(B\_) = -\sum_s \pi(s)P(s)\left[\omega - n - \lambda_2(s)\right]. \tag{54c}$$

As before we will use $\mu(s, B\_) = -\beta V(B(s, B\_))$.

At an interior solution ($\lambda_1(s) = 0, \lambda_2(s) = 0$) labor supply is constant; $L(s) = L^*$ that solves $\omega \Phi^{-1'}\left(\Phi\left(\frac{n\gamma}{1+\gamma}L^*\right)\right) + n = 0$ and is given by,

$$L^* = \left[\frac{n\theta}{\omega - (\omega - n)(1+\gamma)}\right]^{\frac{1+\gamma}{\gamma}}. \tag{55}$$

Under the policy $B(s, B\_) = B\_$, the non negativity constraint (53b) is slack in state s if

$$B\_ \geq \frac{g(s) - \Phi\left(\frac{n\gamma}{1+\gamma}L^*\right)}{\beta^{-1}P(s) - 1}.$$

Define a threshold $\mathcal{B}(\omega)$ as the minimum assets such that a stationary policy for assets, $B(s, B\_) = B\_$ is feasible with $T(s, B\_) = \Phi\left(\frac{n\gamma}{1+\gamma}L^*\right) + B\_(P(s)\beta^{-1} - 1) - g(s) \geq 0$ for all s:

$$\mathcal{B}(\omega) = \max_s \left(\frac{g(s) - \Phi\left(\frac{n\gamma}{1+\gamma}L^*\right)}{\beta^{-1}P(s) - 1}\right). \tag{56}$$

The sign of the $\frac{\partial \mathcal{B}(\omega)}{\partial \omega}$ is the same as the sign of $-\frac{n\gamma}{1+\gamma}\underbrace{\Phi'\left(\frac{n\gamma}{1+\gamma}L^*\right)}_{<0}\underbrace{\frac{\partial L^*}{\partial \omega}}_{>0} > 0$.

For Part 1 we need the following intermediate lemma:

Lemma 4. *If $P \notin \mathcal{P}^*$ and $B\_ < \mathcal{B}(\omega)$, there exists an s such that $B(s, B\_) > B$*

Proof. Suppose not and $B(s, B\_) \leq B\_$ for all s. Since $V(.)$ is concave, $\beta V'(B(s, B\_) \geq \beta V'(B\_)$. Combining equations (54b) with envelope condition (54c) we get $\tilde{\mathbb{E}}V'(B(s, B\_)) \leq V'(B\_)$. Both of these can hold only if for all $V'(B(s, B\_)) = V'(B\_)$ for all s. Non-negativity of $\lambda_2(s)$ and equation (54c) imply that $-\beta V'(B\_) \leq \omega - n$. Using $V'(B(s, B\_)) = V'(B\_)$ and equation (54b) we see that $-\beta V'(B\_) < \omega - n$ implies $\lambda_2(s) > 0$ and $T(s) = 0$ for all s. In absence of transfers, the value function is strictly concave in the neighborhood of $B\_$ and as $P \notin \mathcal{P}^*$, we can use the same steps in Lemma 2 to show that there exists a s such that $B(s, B\_) > B\_$.

Lastly we rule of cases when $-\beta V'(B\_) = \omega - n$. Let $B^{inf} = \inf\{B | -\beta V'(B) = \omega - n \quad \forall s\}$. The budget constraint at B^{inf} is $T(s) + g(s) - \Phi\left(\frac{n\gamma}{1+\gamma}L^*\right) = \frac{P(s)}{\beta}B^{inf} - B(s, B^{inf})$. Concavity of V implies $B(s, B^{inf}) \geq B^{inf}$ for all s. Note that

$$g(s) - \Phi\left(\frac{n\gamma}{1+\gamma}L^*\right) \leq \frac{P(s)}{\beta}B^{inf} - B(s, B^{inf}) \leq B^{inf}\left(\frac{P(s)}{\beta} - 1\right),$$

and from expression (56) $B^{inf} \geq \mathcal{B}(\omega)$. However $\mu(s, B\_) = \omega - n$ and hence by the definition of B^{inf} we have $B\_ \geq B^{inf}$. This contradicts our assumption $B\_ < \mathcal{B}(\omega)$. □

Using Lemma 4 we can construct a sequence that converges to $\mathcal{B}(\omega)$ from and $B_{-1} < \mathcal{B}(\omega)$. To show that B_t converges to $\mathcal{B}(\omega)$ with probability one, we will show $\mathcal{B}(\omega)$ is reached from any $B\_ < \mathcal{B}(\omega)$ with a finite sequence of shocks with probability bounded away from zero. To do this we show that there is a point strictly smaller than $\mathcal{B}(\omega)$ from where $\mathcal{B}(\omega)$ is reached in one step.

Lemma 5. *There exists a $\check{B} < \mathcal{B}(\omega)$ we can find $\bar{s}$ with the property that $B(\bar{s}, B\_) \geq \mathcal{B}(\omega)$ for all $B\_ \geq \check{B}$*

Proof. At $\mathcal{B}(\omega)$, there exist some $\bar{s}$ such that $T(\bar{s}, \mathcal{B}(\omega)) = \epsilon > 0$. Now define $\check{B}$ as follows:

$$\check{B} = \mathcal{B}(\omega) - \frac{\epsilon\beta}{2P(\bar{s})}.$$

Suppose to the contrary $B(\bar{s}, B\_) < \mathcal{B}(\omega)$ for some $B\_ \geq \check{B}$. This implies that $\tau(\bar{s}, B\_) > \tau^*(\omega)$ and $T(\bar{s}, B\_) = 0$.[17] Consider the budget constraint for the government at $B\_$ and $\mathcal{B}(\omega)$ for the shock $\bar{s}$

$$g(\bar{s}) + B(\bar{s}, B\_) = \frac{P(\bar{s})}{\beta} B\_ + n\theta\tau(\bar{s}, B\_)l(\bar{s}, B\_), \tag{57a}$$

$$g(\bar{s}) + \mathcal{B}(\omega) = \frac{P(\bar{s})}{\beta}\mathcal{B}(\omega) + n\theta\tau^*(\omega)l^* - T(\bar{s}, \mathcal{B}(\omega)). \tag{57b}$$

Subtracting equation (57a) from (57b) we get

$$\mathcal{B}(\omega) - B(\bar{s}, B\_) + n\theta\left[\tau(\bar{s}, B\_)l(\bar{s}, B\_) - \tau^*(\omega)l^*\right] = \frac{P(\bar{s})}{\beta}\left(\mathcal{B}(\omega) - B\_\right) - T(\bar{s}, \mathcal{B}(\omega). \tag{58}$$

The left hand side of equation (58) is strictly positive as the first term is greater than zero under our assumption and the second term pertaining to the tax revenues is also strictly positive as the government taxes on the left of the Laffer curve and $\tau(\bar{s}, B\_) > \tau^*(\omega)$. But the right hand side is strictly negative as

$$\frac{P(\bar{s})}{\beta}\left(\mathcal{B}(\omega) - B\_\right) \leq \frac{\epsilon}{2} < \epsilon = T(\bar{s}, \mathcal{B}(\omega)).$$

This yields a contradiction. □

Define $\hat{B}(B\_)$ as $\max_s B(s, B\_)$ and $\hat{s}(B\_)$ as the shock that achieves this maximum. Note that $\hat{B}(B\_) - B\_$ is continuous on $[\underline{B}, \check{B}]$, bounded below by zero by Lemma 4 and therefore attains a minimum at B^{min}. Let $\delta = \hat{B}(B^{min}) - B^{min} > \eta > 0$. Now consider any initial $B\_ \in [\underline{B}, \mathcal{B}(\omega)]$. If $B\_ \geq \check{B}$, then by Lemma 5 we know that $\mathcal{B}(\omega)$ will be reached in one shock. Otherwise if $B\_ < \check{B}$, we can construct a sequence of shocks $s_t = \hat{s}(B_{t-1})$ of length $N = \frac{\check{B} - \underline{B}}{\delta}$. There exits $t < N$ such that $B_t > \check{B}$. We can use the same arguments as in the last part of the proof of Proposition 2 to conclude that $\Pr\{B_t \to \mathcal{B}(\omega)\} = 1$.

When $P \in \mathcal{P}^*$ such that $B^*(P) < \mathcal{B}(\omega)$ then there are two different cases: $B_{-1} \leq B^*$ and $B^* > B_{-1} > \mathcal{B}(\omega)$. When $B_{-1} < B^*$ we can guess that the optimal policy satisfies $T_t = 0$ for all t. Under this guess, it can be verified that the same policy rules that solve (36) satisfy the first order conditions for problem (52). We can therefore conclude that $\mu_t \to \mu^*$and $B_t \to B^*$. For $B^* > B_{-1} > \mathcal{B}(\omega)$, we will order $\mu(s, B\_)$ relative to $P(s)$ for $B\_ > B^*$ such that $\mu(s', B\_) \geq \mu(s'', B\_)$ when $P(s') \geq P(s'')$. To see this note that if $T(s') > 0$ then $\mu(s', B\_) = \omega - n \geq \mu(s'', B\_)$. For $T(s') = 0$

$$g(s') - \frac{1}{\beta}P(s')B\_ \leq g(s'') - \frac{1}{\beta}P(s'')B\_ \leq g(s'') - \frac{1}{\beta}P(s'')B\_ + T(s'').$$

[17] This follows from the fact that $B(\bar{s}, B\_) < \mathcal{B}(\omega)$ implies $\mu(\bar{s}, B\_) < \mu(\mathcal{B}(\omega))$ and (54a).

Equation (54b) gives us

$$\omega\Phi^{-1'}\left(B(s'')-\frac{1}{\beta}P(s'')B\_+g(s'')+T(s'')\right)+\beta V'(B(s''))=\omega\Phi^{-1'}\left(B(s')-\frac{1}{\beta}P(s')B\_+g(s')\right)+\beta V'(B(s')).$$

Concavity of Φ^{-1} and V implies $B(s',B\_)\geq B(s'',B\_)$ and $\mu(s',B\_)\geq\mu(s'',B\_)$ as desired. The ordering implies that $\text{cov}(\mu(s,B\_),P(s))\geq 0$ for $B^*>B_{-1}>\mathcal{B}(\omega)$ and hence $\mu_t\geq\mathbb{E}\mu_{t+1}$. From the property that $\mu(s,B^*)=\mu^*$, monotonicity, and continuity of function $\mu(s,B\_)$ with respect to $B\_$ we have $\mu_t\geq\mu^*$ for all $t\geq 0$. Standard martingale convergence arguments imply that μ_t converges almost surely. As there are only two steady states: $\mu^*(P^*)$ and $\omega-n$, B_t will converge to one of the two associated steady state levels of a government assets: B^* or $\mathcal{B}(\omega)$. $\square$

7.6 Bellman equation for $t=0$

Let $V_0\left(\{\tilde{b}_{i,-1}\}_{i=1}^{I-1},s_0\right)$ be the value to the planner at $t=0$, where $\tilde{b}_{i,-1}$ denotes initial debt. We retain the normalization $R_0=\beta^{-1}P(s_0)$. The $t=0$ Bellman equation solves,

$$V_0\left(\{\tilde{b}_{i,-1}\}_{i=1}^{I-1},s_0\right)=\max_{a_0,x_0,\rho_0}\sum_i\omega_iU^i(c_{i,0},l_{i,0})+\beta V\left(x_0,\rho_0,s_0\right)$$

subject to,

$$U^I_{c,0}\left[c_{i,0}-c_{I,0}\right]+x_{i,0}+\left(U^i_{l,0}l_{i,0}\frac{U^I_{c,0}}{U^i_{c,0}}-U^I_{l,0}l_{I,0}\right)=\beta^{-1}P(s_0)U^I_{c,0}\tilde{b}_{i,-1}\text{ for all }i<I$$

$$\frac{U^i_{l,0}}{\theta_{i,0}U^i_{c,0}}=\frac{U^I_{l,0}}{\theta_{I,0}U^{I,0}_c}\text{ for all }i<I$$

$$\sum_i n_ic_{i,0}+g_0=\sum_i n_i\theta_{i,0}l_{i,0}$$

$$\rho_{i,0}=\frac{U^i_{c,0}}{U^I_{c,0}}\ \forall\ i<I$$

7.7 Formula for optimal taxes

The first order conditions with respect to $l_i(s)$ and $l_I(s)$ are,

$$\begin{aligned}\omega_iU^i_l(s)-\frac{1}{\rho_i'(s)}\mu_i(s)\left[U^i_{ll}(s)l_i(s)+U^i_l(s)\right]-\frac{\phi_i(s)U^i_{ll}(s)}{\theta_i(s)U^i_c(s)}+n_i\theta_i(s)\xi(s)=0\\ \omega_IU^I_l(s)+U^I_c(s)\left(\sum_{i<I}\mu_i(s)\right)\left[U^I_{ll}(s)l_I(s)+U^I_l(s)\right]+\frac{U^I_{ll}(s)}{\theta_I(s)U^I_c(s)}\left(\sum_{i<I}\phi_i(s)\right)+n_I\theta_I(s)\xi(s)=0\end{aligned}\tag{60}$$

Multiplying each equation by $\frac{\theta^i(s)U^i_c(s)}{U^i_{ll}(s)}$ and using $\frac{l_{i,}(s)U^i_{ll}(s)}{U^i_l(s)}=\gamma$ we recover,

$$\omega_i \frac{U_c^i(s) y_i(s)}{\overline{\gamma}} - U_c^I(s)\mu_i(s) y_i(s)\left[1+\frac{1}{\overline{\gamma}}\right] - \phi_i(s) + n_i y_i(s)\xi(s)\frac{\theta_i(s)U_c^i(s)}{\overline{\gamma}U_l^i(s)} = 0$$

$$\omega_I \frac{U_c^I(s) y_I(s)}{\overline{\gamma}} + U_c^I(s)\left(\sum_{i<I}\mu_i(s)\right) y_I(s)\left[1+\frac{1}{\overline{\gamma}}\right] + \sum_{i<I}\phi_i(s) + n_I y_I(s)\xi(s)\frac{\theta_I(s)U_c^I(s)}{\overline{\gamma}U_l^I(s)} = 0.$$

Define $\hat{\omega}_i = \frac{\omega_i}{n_i}$, $\hat{\mu}_i = \frac{U_c^I \mu_i - U_c^i \omega_i}{n_i}$ for $i < I$ and $\hat{\mu}_I = -\frac{U_c^I \sum_{i<I}\mu_i}{n_I} - n_I^{-1} U_c^I \omega_I$. The equations above can be combined to get

$$\frac{1}{1-\tau(s)} = -\left(\frac{\sum_i n_i \left[\gamma \hat{\omega}_i U_c^i(s) + (1+\gamma)\hat{\mu}_i(s)\right] y_i(s)}{\xi(s)\sum_i n_i y_i(s)}\right). \tag{61}$$

Formula (25) follows from applying $\sum_i n_i w_i(s) y_i(s) = \left(\sum_i n_i w_i(s)\right)\left(\sum_i n_i y_i(s)\right) + \sum_i n_i \left(w_i(s) - \sum_i n_i w_i(s)\right)\left(y_i(s) - \sum_i n_i y_i(s)\right) = \overline{wy} + \operatorname{cov}(w_i(s), y_i(s))$.

We next specialize (61) to the two special cases. For the quasilinear case, at an interior solution, the first order condition with respect to $c_i(s)$ gives us $\hat{\mu}_i(s) = -\xi(s)$. Furthermore, by construction, $\sum_i n_i \hat{\mu}_i(s) = -\sum \omega_i = -1$ and substituting in equation (61) we obtain

$$\frac{1}{1-\tau(s)} = 1 - \gamma\left(\frac{\sum_i n_i \hat{\omega}_i y_i(s)}{\bar{y}(s)} - 1\right).$$

For the case without savings, let $\tilde{c}_i = c_i - T$ be net of transfers consumption for agent i. In absence of (24b) and (24e), the first order condition with respect to $\tilde{c}_i$ can be expressed as

$$-\hat{\mu}_i n_i + \phi_i n_i (1-\tau)\theta_i U_{cc}^i - n_i \xi = 0, \tag{62}$$

Eliminate ϕ_i from equations (60) , substituting $\sigma_i = -U_{cc}^i \tilde{c}_i / U_c^i$ and collecting terms we get,

$$n_i \hat{\mu}_i = -n_i \xi \left(\gamma + \frac{\sigma_i}{(1-\tau)}\right)\frac{1}{\gamma + \sigma_i}. \tag{63}$$

Using this expression for $\hat{\mu}_i$ in equation (61) and rearranging terms gives us Formula (27).

7.8 Steady states for more general economies

Restrict shocks to be i.i.d. For constructing steady states we begin with a choice for $\{\tau(s), \boldsymbol{\rho}(s)\}$ This determines a competitive equilibrium allocation $\{c_i(s), l_i(s)\}_i$ using equations (24c), (24d) and (24e). Let $Z_i\left(s|\{\tau(s),\boldsymbol{\rho}(s)\}\right) = U_c^I(s)\left[c_i(s) - c_I(s)\right] + \left(\frac{U_l^i(s)}{U_c^i(s)} U_c^I(s) - l_I(s) U_l^I(s)\right)$, the implementability constraints for the planner reduces to

$$Z_i(s) + x_i'(s) = \frac{\mathcal{R}(s)}{\beta} x_i \text{ for all } s, i < I, \tag{64a}$$

$$\mathbb{E}\mathcal{R}\rho_i' = \rho_i \text{ for } i < I. \tag{64b}$$

Let $\pi(s)\mu_i(s)$ and λ_i be Lagrange multipliers on constraints (64a) and (64b). Imposing the restrictions $x_i'(s) = x_i$ and $\rho_i'(s) = \rho_i$, at a steady state $\{\mu_i, \lambda_i, x_i, \rho_i\}_{i=1}^{I-1}$ and $\{\tau(s)\}_s$ are determined by the following equations:

$$Z_i(s) + x_i'(s) = \frac{\mathcal{R}(s)}{\beta} x_i \text{ for all } s, i < I, \tag{65a}$$

$$U_\tau(\tau(s), \boldsymbol{\rho}, s) - \sum_i \mu_i Z_{i,\tau}(\tau(s), \boldsymbol{\rho}, s) = 0 \text{ for all } s, \tag{65b}$$

$$U_{\rho_i}(\tau(s), \boldsymbol{\rho}, s) - \sum_j \mu_j Z_{j,\rho_i}(\tau(s), \boldsymbol{\rho}, s) + \lambda_i(\mathcal{R}(s) + \beta) = 0. \text{ for all } s, i < I \tag{65c}$$

When the shock s takes only two values, (65) is a square system of $4(I-1)+2$ equations in $4(I-1)+2$ unknowns $\{\mu_i^{SS}, \lambda_i^{SS}, x_i^{SS}, \rho_i^{SS}\}_{i=1}^{I-1}$ and $\{\tau^{SS}(s)\}_s$.

7.9 Proof of Proposition 6

Proof. With log quadratic preferences and i.i.d shocks we can reduce (65) to

$$x(s) = \frac{1 + \rho[l_1(\rho, s)^2 - 1]}{\frac{P(s)/c_2(\rho,s)}{\beta \mathbb{E}[\frac{P}{c_2}](\rho)} - 1}. \tag{66}$$

The functions $l_1(\rho, s)$ and $c_2(\rho, s)$ are given by

$$l_1(s) = \frac{g(s) + \sqrt{g(s)^2 + 4C(\rho)\theta_1^2}}{2\theta_1}, \tag{67}$$

$$\frac{1}{c_2(s)} = \left(\frac{1+\rho}{C(\rho)}\right)\left(\frac{g(s) + \sqrt{g(s)^2 + 4C(\rho)\theta_1^2}}{2\theta_1^2}\right),$$

where $C(\rho) = \frac{1+\rho}{\omega_1(1-\rho)+2\rho^2\omega_2}$. Since both these functions are increasing in g, they are ordered with $s \in \{s_H, s_L\}$ and $l_{1,l} < l_{1,h}$ and $\frac{1}{c_{2,l}} < \frac{1}{c_{2.h}}$.

Existence of a steady state is equivalent to finding a scalar ρ that solves $x(s_l) = x(s_h)$. When $P(s) = 1$ this amounts to showing that the following function crosses zero:

$$f(\rho) = \frac{1 + \rho[l_1(\rho, s_l)^2 - 1]}{1 + \rho[l_1(\rho, s_h)^2 - 1]} - \frac{\frac{1/c_2(\rho,s_l)}{\beta \mathbb{E}[\frac{P}{c_2}](\rho)} - 1}{\frac{1/c_2(\rho,s_h)}{\beta \mathbb{E}[\frac{P}{c_2}](\rho)} - 1}.$$

The main step is to show that $f(.)$ takes both negative and positive values and then we can appeal to the intermediate value theorem.

Ordering of l_1 and c_2 imply that for all $\rho > 0$

$$1 + \rho[l_1(\rho, s_h)^2 - 1] > 1 + \rho[l_1(\rho, s_l)^2 - 1], \tag{68a}$$

$$\frac{1/c_2(\rho, s_h)}{\beta \mathbb{E}[\frac{P}{c_2}](\rho)} - 1 > \frac{1/c_2(\rho, s_l)}{\beta \mathbb{E}[\frac{P}{c_2}](\rho)} - 1. \tag{68b}$$

Using equation (67), we can construct a $\underline{\rho}$ [18] such that approaching $\underline{\rho}$ from the right yields the following limits,

$$\lim_{\rho\to\underline{\rho}^+} \frac{1+\rho[l_1(\rho,s_h)^2-1]}{1+\rho[l_1(\rho,s_l)^2-1]} = 1,$$

$$\lim_{\rho\to\underline{\rho}^+} \frac{\frac{1/c_2(\rho,s_l)}{\beta\mathbb{E}[\frac{P}{c_2}](\rho)}-1}{\frac{1/c_2(\rho,s_h)}{\beta\mathbb{E}[\frac{P}{c_2}](\rho)}-1} < 1.$$

This implies that $\lim_{\rho\to\underline{\rho}^+} f(\underline{\rho}) > 0$. Next taking the limit as $\rho\to\infty$ we see that $C(\rho)\to 0$ which, with $\frac{g(s)}{\theta_1}<1$, implies

$$\lim_{\rho\to\infty} 1+\rho[l_1(\rho,s)^2-1] = -\infty.$$

Consequently there exists $\overline{\rho}$ such that $1+\overline{\rho}[l_1(\overline{\rho},s_h)^2-1]=0$.[19] From equation (68a),

$$0 = 1+\overline{\rho}[l_1(\overline{\rho},s_h)^2-1] > 1+\overline{\rho}[l_1(\overline{\rho},s_l)^2-1],$$

which implies that $\lim_{\rho\to\overline{\rho}^-}\frac{1+\rho[l_1(\rho,s_l)^2-1]}{1+\rho[l_1(\rho,s_h)^2-1]} = -\infty$ and along with $\frac{\frac{1/c_2(\rho,s_l)}{\beta\mathbb{E}[\frac{P}{c_2}]}-1}{\frac{1/c_2(\rho,s_h)}{\beta\mathbb{E}[\frac{P}{c_2}]}-1} \geq -1$, we get $\lim_{\rho\to\overline{\rho}^-} f(\rho)=-\infty$. Thus there exists ρ_{SS} such that $f(\rho_{SS})=0$.

Finally, as $\rho_{SS}<\overline{\rho}$ we know that $1+\rho_{SS}[l_1(\rho_{SS},s_h)-1]>0$ and as $\frac{1/c_2(\rho,s_h)}{\beta\mathbb{E}[\frac{P}{c_2}]}>1$, we conclude,

$$x_{SS} = \frac{1+\rho_{SS}[l_1(\rho_{SS},s_h)-1]}{\frac{1/c_2(\rho,s_h)}{\beta\mathbb{E}[\frac{P}{c_2}](\rho)}-1} > 0.$$

The ordering for $\mathcal{R}(s)$ follows from (68b).

For the second part of the statement in Proposition 6, we need to construct a steady steady state with $x_{SS}<0$. Let ρ_{SS} be such that $0>1+\rho_{SS}[l_1(\rho_{SS},s_h)^2-1]>1+\rho_{SS}[l_1(\rho_{SS},s_l)^2-1]$. Choose a payoff vector P such that $1<\frac{P(s)/c_2(\rho_{SS},s)}{\beta\mathbb{E}[\frac{P}{c_2}]}$ for all s and

$$1 > \frac{\frac{P(s_h)/c_2(\rho_{SS},s_h)}{\beta\mathbb{E}[\frac{P}{c_2}]}-1}{\frac{P(s_l)/c_2(\rho_{SS},s_l)}{\beta\mathbb{E}[\frac{P}{c_2}]}-1} = \frac{1+\rho_{SS}[l_1(\rho_{SS},s_h)^2-1]}{1+\rho_{SS}[l_1(\rho_{SS},s_l)^2-1]}. \tag{69}$$

The steady state x_{SS} solves

$$x_{SS} = \frac{1+\rho_{SS}[l_1(\rho_{SS},s_h)^2-1]}{\frac{P(s_l)/c_2(\rho_{SS},s_h)}{\beta\mathbb{E}[\frac{P}{c_2}]}-1} < 0.$$

At such a P, the ordering for $\mathcal{R}(s)$ follows from equation (69).

□

[18] Note that we will have to deal with two different cases. If $\omega_1(1-\rho)+2\rho^2\omega_2>0$ for all $\rho\geq 0$ then $C(\rho)\geq 0$ for all $\rho\geq 0$ and is bounded above and $\underline{\rho}=0$. If $\omega_1(1-\rho)+2\rho^2\omega_2=0$ for some $\rho>0$ let $\underline{\rho}$ be the largest positive root of $\omega_1(1-\rho)+2\rho^2\omega_2$. Note that $\lim_{\rho\to\underline{\rho}^+}C(\rho)=\infty$. In the first case $\underline{\rho}=0$ and in the second case $l_1(\rho,s_l)=l_1(\rho,s_h)$ as $\rho\to\underline{\rho}^+$.

[19] This can be seen from the fact $\lim_{\rho\to\underline{\rho}^+}1+\rho[l_1(\rho,s_l)^2-1]>0$ and $\lim_{\rho\to\infty}1+\rho[l_1(\rho,s_l)^2-1]>-\infty$, thus $\overline{\rho}$ exists in $(\underline{\rho},\infty)$

7.10 Numerical approximation to a Ramsey plan

This appendix applies a method of Evans (2014) to approximate a Ramsey plan for the I type of agents economy of section 5. To focus on essential steps, we take a special case where productivities for agent i are $\log\theta_{i,t} = \log\overline{\theta}_i + \sigma_\epsilon\epsilon_t$ and where ϵ_t is i.i.d with mean zero and variance one.

We begin by re-writing the the Bellman equation (23) in section 4 in way that makes it convenient to apply the numerical algorithm. Let $x\_ = U^i_{c\_}b_i$ and $m^i\_ \propto \frac{1}{U^i_c}$ with $\sum_i m^i = 1$. [20]The modified Bellman equation for $t \geq 1$ is:

$$V(\boldsymbol{x}_-, \boldsymbol{m}_-) = \max \sum_s \left[\Pi(s)\sum_i \omega^i U(c^i(s), l^i(s)) + \beta V(\boldsymbol{x}(s)\boldsymbol{m}(s))\right] \tag{70}$$

subject to

$$\begin{aligned}
\frac{\pi(s)U^i_c(s)x^i_-}{\beta\mathbb{E}_- P U^i_c} &= U^i_c(s)(c^i(s) - T(s)) + U^i_l(s)l^i(s) + x^i(s)\\
U^i_c(s)\exp\epsilon(s)\theta^i(1-\tau_l(s)) &= -U^i_l(s)\\
\alpha(s) &= m^i(s)U^i_c(s)\\
\gamma_- &= m^i_-\mathbb{E}_- P U^i_c\\
\sum_i n_i\left[\exp\epsilon(s)\theta^i l^i(s) - c^i(s)\right] &= 0\\
\sum_i n^i\frac{x^i(s)}{m^i(s)} &= 0\\
\sum_i n^i m^i(s) &= 1
\end{aligned}$$

It is easy to check that satisfying (71) is equivalent to satisfying (24a)-(24e). For what follows next, we use the notation,

- z_i, as individual state variables that include m_i, μ_i.
- y_i, as individual's choice variables that include m_i, μ_i, c_i, l_i and Lagrange multipliers on individual constraints.
- Y, as the planner's aggregate choice variables that include $\tau_l, T, \alpha, \gamma_-$ and Lagrange multipliers on aggregate constraints.

Next, stack the first order conditions of problem (70) and constraints (71) in the following form,

$$\mathbb{E}_{t-1}F(y_{i,t}, \mathbb{E}_{t-1}y_{i,t}, Y_t, y_{i,t+1}, \epsilon_t; \boldsymbol{z}_{t-1}, \sigma_\epsilon) = 0, \tag{72a}$$

$$\sum_i n_i G(y_{i,t}, Y, \epsilon_t; \boldsymbol{z}_{t-1}, \sigma_\epsilon) = 0. \tag{72b}$$

[20]Note that Ricardian equivalence implies that we can normalize $\sum_i n^i \frac{x^i_-}{U^i_{c\_}} = 0$. Thus, the dimension of the state variables appearing in our modified Bellman equation is also $2I-2$, as in problem (23). However, by not normalizing with respect to some arbitrary agent's asset holdings (such as $i=1$), we attain a symmetry that turns out to be convenient.

The *Ramsey plan* can then be represented as a set of functions (y, Y, Z) that solve (72):

$$
\begin{aligned}
y_{i,t} &= y_i(\epsilon_t, \boldsymbol{z}_{t-1}, \sigma_\epsilon), \quad i = 1, \dots I \\
Y_t &= Y(\epsilon_t; \boldsymbol{z}_{t-1}, \sigma_\epsilon) \\
\boldsymbol{z}_t &= Z(\epsilon_t; \boldsymbol{z}_{t-1})
\end{aligned}
$$

Our goal is to approximate outcomes generated by the system of equations (73). To do this, we generate a sequence of approximations to the system of equations (73). We generate the t^{th} outcome along a sample path by drawing ϵ_t and applying our approximation of equations (73) for date t. Our approximation to these functions at date t depends on the outcomes $\boldsymbol{z}_{t-1}$ generated at the previous step of the simulation. To generate functions approximating (73) at date t, we use a small-noise expansion (i.e., around $\sigma_\epsilon = 0$) to those functions at state $\boldsymbol{z}_{t-1}$, an expansion that exploits economic properties associated with the limiting $\sigma_\epsilon = 0$ economy at that state. Thus, to approximate sample paths drawn from the recursive system (73), we use a sequence of Taylor series approximations around a sequence of points generated endogenously during a simulation.

The steps of the algorithm proceed sequentially as follows:

1. Given some $\boldsymbol{z}_{t-1}$, compute the individual and aggregate choice variables in a limiting economy with $\sigma_\epsilon = 0$. For our problem, we choose state variables that ensure that $\boldsymbol{z}_t = \boldsymbol{z}_{t-1}$ in this limiting economy.[21] The allocation in the limiting economy is a set of values $\left(\{\bar{y}_i, \}_i, \bar{Y}\right)$ that solve (72a) and (72b) at $\sigma_\epsilon = 0$. This logic gives us a set of non-linear equations

$$F(\bar{y}_i, \bar{y}_i, \bar{Y}, \bar{y}_i, 0; \bar{z}_{t-1}, 0) = 0 \quad \forall i \tag{74a}$$

$$\sum_i n_i G(\bar{y}, \bar{Y}, 0; \boldsymbol{z}_{t-1}, \sigma_\epsilon) = 0 \tag{74b}$$

   whose solution $(\bar{y}, \bar{Y})$ depends on $\boldsymbol{z}_{t-1}$. It is significant that the "steady state" $(\bar{y}, \bar{Y})$ would be the outcome for a complete markets economy with initial condition $\boldsymbol{z}_{t-1}$. This follows partly from the fact that when $\sigma_\epsilon = 0$, a risk free bond is enough to complete markets.

2. Next construct a truncated Taylor series approximation to the functions y, Y appearing in (73) around the steady state $\bar{y}(\boldsymbol{z}_{t-1})$ and $\bar{Y}(\boldsymbol{z}_{t-1})$ obtained in the step 1; this yields approximations

$$
\begin{aligned}
y_i(\epsilon_t; \boldsymbol{z}_{t-1}, \sigma_\epsilon) \approx \bar{y}_i(\boldsymbol{z}_{t-1}) &+ \frac{\partial y_i}{\partial \epsilon}(0; \boldsymbol{z}_{t-1}, 0)\epsilon_t \\
&+ \frac{1}{2}\frac{\partial^2 y_i}{\partial \epsilon^2}(0; \boldsymbol{z}_{t-1}, 0)\epsilon_t^2 \\
&+ \frac{1}{2}\frac{\partial^2 y_i}{\partial \sigma_\epsilon^2}(0; \boldsymbol{z}_{t-1}, 0)\sigma_\epsilon^2
\end{aligned} \tag{75a}
$$

[21] Extensions to more general environments where there do not exist such steady states or where Γ_t follows a deterministic path in the non stochastic limit can be found in Evans (2014).

and

$$Y(\epsilon_t; \boldsymbol{z}_{t-1}, \sigma_\epsilon) \approx \bar{Y}(\boldsymbol{z}_{t-1}) + \frac{\partial Y}{\partial \epsilon}(0; \boldsymbol{z}_{t-1}, 0)\epsilon_t + \frac{1}{2}\frac{\partial^2 Y}{\partial \epsilon_t^2}(0; \boldsymbol{z}_{t-1}, 0)\epsilon_t^2 + \frac{1}{2}\frac{\partial^2 Y}{\partial \sigma_\epsilon^2}(0; \boldsymbol{z}_{t-1}, 0)\sigma_\epsilon^2. \tag{75b}$$

The main computational task is to evaluate derivatives at the steady state. This involves totally differentiating system (72a) and (72b) at the non stochastic steady state associated with $\boldsymbol{z}_{t-1}$. [22]

3. Draw shocks ϵ_t and use the approximate policies in (75a) and (75b) to obtain $y_{i,t}$ and Y_t. Remember that $z_{i,t}$ is assumed to be included in the vector $y_{i,t}$. This yields us the next $\boldsymbol{z}_t$.

4. Advance to $t+1$ and repeat steps 1 to 3 using the updated $\boldsymbol{z}_t$ as the initial distribution

References

Aguiar, Mark, and Manuel Amador. 2011. "Growth in the Shadow of Expropriation." *The Quarterly Journal of Economics*, 126(2): 651–697, URL: `http://ideas.repec.org/a/oup/qjecon/v126y2011i2p651-697.html`. 1.1

Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala. 2002. "Optimal Taxation without State âĂŘ Contingent Debt." *Journal of Political Economy*, 110(6): 1220–1254, URL: `http://www.jstor.org/stable/10.1086/343744`. 1, 1.1, 3, 3, 8, 6

Angeletos, George-Marios. 2002. "Fiscal Policy With Noncontingent Debt And The Optimal Maturity Structure." *The Quarterly Journal of Economics*, 117(3): 1105–1131. 1.1

Atkinson, Anthony Barnes, and Joseph E Stiglitz. 1976. "The design of tax structure: direct versus indirect taxation." *Journal of public Economics*, 6(1): 55–75. 4.1

Azzimonti, Marina, Eva de Francisco, and Per Krusell. 2008a. "Aggregation and Aggregation." *Journal of the European Economic Association*, 6(2-3): 381–394, URL: `http://doi.wiley.com/10.1162/JEEA.2008.6.2-3.381`, DOI: `http://dx.doi.org/10.1162/JEEA.2008.6.2-3.381`. 4

Azzimonti, Marina, Eva de Francisco, and Per Krusell. 2008b. "Production subsidies and redistribution." *Journal of Economic Theory*, 142(1): 73–99, URL: `http://linkinghub.elsevier.com/retrieve/pii/S0022053107001020`, DOI: `http://dx.doi.org/10.1016/j.jet.2007.03.009`. 4

Barro, Robert J. 1974. "Are government bonds net wealth?." *The Journal of Political Economy*, 82(6): 1095–1117, URL: `http://www.jstor.org/stable/10.2307/1830663`. 1.1

[22] For large I, calculating these derivatives can be further simplified for a class of problems where $\frac{\partial z_{i,t}}{\partial z_{i,t-1}}(0; z_{,t-1}, 0)$ is independent of $z_{i,t-1}$. In our context this turns out to be an identity matrix.

Barro, Robert J. 1979. "On the determination of the public debt." *The Journal of Political Economy*, 87(5): 940–971, URL: http://www.jstor.org/stable/10.2307/1833077. 1.1, 6

Bassetto, Marco. 1999. "Optimal Fiscal Policy with Heterogeneous Agents." 1.1

Bassetto, Marco, and Narayana Kocherlakota. 2004. "On the irrelevance of government debt when taxes are distortionary." *Journal of Monetary Economics*, 51(2): 299–304, URL: http://linkinghub.elsevier.com/retrieve/pii/S0304393203001430, DOI: http://dx.doi.org/10.1016/j.jmoneco.2002.12.001. 2

Buera, Francisco, and Juan Pablo Nicolini. 2004. "Optimal maturity of government debt without state contingent bonds." *Journal of Monetary Economics*, 51(3): 531–554. 1.1

Chari, V V, Lawrence J Christiano, and Patrick J Kehoe. 1994. "Optimal Fiscal Policy in a Business Cycle Model." *Journal of Political Economy*, 102(4): 617–652, URL: http://www.nber.org/papers/w4490http://www.jstor.org/stable/2138759, DOI: http://dx.doi.org/10.2307/2138759. 1.1, 4, 6

Correia, Isabel. 2010. "Consumption taxes and redistribution." *American Economic Review*, 100(September): 1673–1694, URL: http://www.ingentaconnect.com/content/aea/aer/2010/00000100/00000004/art00014. 4

Diamond, Peter A. 1975. "A many-person Ramsey tax rule." *Journal of Public Economics*, 4(4): 335–342. 4.1

Evans, David. 2014. "Perturbation Theory with Heterogeneous Agents." 1, 5, 7.10, 21

Faraglia, Elisa, Albert Marcet, and Andrew Scott. 2012. "Dealing with Maturity : Optimal Fiscal Policy with Long Bonds." 1

Farhi, Emmanuel. 2010. "Capital Taxation and Ownership When Markets Are Incomplete." *Journal of Political Economy*, 118(5): 908–948, URL: http://www.jstor.org/stable/10.1086/657996. 1, 4, 6

Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning. 2007. "New dynamic public finance: a user's guide." In *NBER Macroeconomics Annual 2006, Volume 21*. 21: MIT Press, 317–388, URL: http://www.nber.org/chapters/c11181.pdf. 2

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. "The Nature of Countercyclical Income Risk." *Journal of Political Economy*, 122(3): pp. 621–660, URL: http://www.jstor.org/stable/10.1086/675535. 5.1

Hopenhayn, Hugo A., and Edward C. Prescott. 1992. "Stochastic Monotonicity and Stationary Distributions for Dynamic Economies." *Econometrica*, 60(6): pp. 1387–1406, URL: http://www.jstor.org/stable/2951526. 7.3

Kydland, Finn E, and Edward C Prescott. 1980. "Dynamic optimal taxation, rational expectations and optimal control." *Journal of Economic Dynamics and Control*, 2(0): 79–91, URL: http://www.sciencedirect.com/science/article/pii/0165188980900524, DOI: http://dx.doi.org/http://dx.doi.org/10.1016/0165-1889(80)90052-4. 4

Lucas, Robert E, and Nancy L Stokey. 1983. "Optimal fiscal and monetary policy in an economy without capital." *Journal of Monetary Economics*, 12(1): 55–93, URL: http://www.sciencedirect.com/science/article/pii/0304393283900491, DOI: http://dx.doi.org/http://dx.doi.org/10.1016/0304-3932(83)90049-1. 1.1, 3, 3, 6

McKay, Alisdair, and Ricardo Reis. 2013. "The role of automatic stabilizers in the US business cycle."Technical report, National Bureau of Economic Research. 1.1

Newcomb, Simon. 1865. *A critical examination of our financial policy during the Southern rebellion.* (document), 2.1

Shin, Yongseok. 2006. "Ramsey meets Bewley: Optimal government financing with incomplete markets." *Unpublished manuscript, Washington University in St. Louis*(August): . 1.1

Shin, Yongseok. 2007. "Managing the maturity structure of government debt." *Journal of Monetary Economics*, 54(6): 1565–1571, URL: http://ideas.repec.org/a/eee/moneco/v54y2007i6p1565-1571.html. 1.1

Werning, Ivan. 2007. "Optimal Fiscal Policy with Redistribution,." *Quarterly Journal of Economics*, 122(August): 925–967, URL: http://qje.oxfordjournals.org/content/122/3/925.abstract, DOI: http://dx.doi.org/10.1162/qjec.122.3.925. 1.1, 4.2, 11

Werning, Ivan. 2012. "Notes on Tax Smoothing with Heterogeneous Agents." 3

Woodford, Michael. 1990. "Public Debt as Private Liquidity." *American Economic Review*, 80(2): 382–88, URL: http://ideas.repec.org/a/aea/aecrev/v80y1990i2p382-88.html. 6