Taxation, debt, and redistribution*

Anmol Bhandari, David Evans, Mikhail Golosov, Thomas Sargent

September 2013

Abstract

We study optimal income taxes and transfers in an economy with heterogeneous agents and aggregate shocks. The net distribution of debt holdings across agents influences optimal allocations, transfers, and tax rates, but the level of government debt does not. Higher cross-section correlations of debt holdings and labor incomes imply more distortions and lower welfare. In incomplete markets economies, setting taxes and transfers optimally substantially alters the character of the government's precautionary incentive to accumulate assets relative to representative agent Ramsey models like Aiyagari et al. (2002). We analyze how the government's long-run asset or debt position emerges from possibly countervailing incentives that confront the Ramsey planner. Its distributional motives make the Ramsey planner want to smooth transfers and also possibly to manipulate the correlation between the interest rate and the government's net asset position vis a vis high skilled workers. Related forces also shape higher frequency optimal government responses to recessions accompanied by shifts in the cross-section distribution of skills, responses that differ markedly from those emerging from models with a representative consumer.

Key words: Finding the state is an art. Distorting taxes. Transfers. Redistribution. Government debt. Interest rate risk.

*We thank Mark Aguiar, Stefania Albanesi, Manuel Amador, Andrew Atkeson, Marco Bassetto, V.V. Chari, Harold L. Cole, Guy Laroque, Robert E. Lucas, Jr., Ali Shourideh, Pierre Yared and seminar participants at Bocconi, Chicago, EIEF, the Federal Reserve Bank of Minneapolis, IES, Princeton, Stanford, UCL, Universidade Católica, 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.... if it were 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

This paper studies an economy with agents who differ in their productivities and a benevolent government that imposes an affine income tax that consists of a distortionary proportional tax on labor income and a lump-sum tax or transfer. Figure 1 shows that an affine structure better approximates the US tax-transfers system than just proportional labor taxes. We impose no restrictions on the sign of transfers. If some agents are sufficiently poor or if the government wants enough redistribution, the government always chooses positive transfers. For most of the paper, we study an economy without capital in which a one-period risk-free bond is the only financial asset traded.

In this economy, concerns for redistribution impart a welfare cost of fluctuating transfers, which affects the optimal government response. A decrease in transfers in response to adverse aggregate shocks disproportionately affects agents with low present value of earning. We show that welfare cost of such decrease depends on the difference in asset holdings across households with different productivities, which we call net assets. Our first set of results extends a Ricardian equivalence type of reasoning to argue that gross asset positions (in particular the level of government debt) do not affect the set of feasible allocations that can be implemented in competitive equilibria with taxes and transfers. For example, an increase in government debt which is shared equally by all agents and thus leaves net assets unchanged, has no welfare consequences, in line with the principles proclaimed by Simon Newcomb (1865) in the quotation with which we began this paper. This result also implies (a) that Ricardian equivalence holds in the presence of distorting taxes; (b) that ad-hoc borrowing limits do not restrict the government's ability to respond to shocks; This logic generalizes to structures with complete or incomplete asset markets and more general structures of taxes, with and without physical capital.

The second set of results characterizes the long run behavior of the distribution of assets, taxes and transfers. In a special case of our economy when the aggregate shocks are iid and take two values, we show that generally there exists a steady state in which marginal utility-adjusted net asset positions are constant and taxes and transfers depend only on current realization of shocks. We also identify conditions under which this steady state is stable. In the steady state the magnitude of fluctuations in taxes is low for many standard preferences and zero when preferences are CES. For more general shock processes we show numerically that while a steady state does not exists, similar principles apply - there is an ergodic set to which net asset position converge and in that region the fluctuations of taxes and transfers are diminished.

We identify two forces that determine the properties of the ergodic set. The first force comes from the desire to minimize the welfare cost of fluctuating transfers, for which the Ramsey planner has incentives to sets taxes and transfer that generate negative correlation between households productivities and net assets. Properly recognizing the relevance of net but not gross asset positions stressed in our first set of findings, this can be accomplished by having the government effectively accumulate risk-free claims on high-skilled workers. But another, possibly countervailing or possibly reinforcing, force comes from the Ramsey planner's incentive to use fluctuations in the interest rate to compensate for missing asset markets. If interest rates are high (low) when revenue needs are high, it creates an incentive for the government to accumulate assets (debt) vis-a-vis high-skilled worker. Thus, depending on the comovement of the interest rate with other fundamentals, these two forces may either reinforce each other (for example, in response to a pure TFP shock where the implied interest rates are countercyclical) or go in opposite direction (for example, in the case where interest rates are procylical when TFP shocks are negatively correlated discount factor shocks).

The implication for the optimal policy is particularly stark when agents have quasi-linear preferences and aggregate shocks affect only exogenous government expenditures. In this case both of the forces are absent and for any initial asset distribution economy is immediately in the steady state in which assets and taxes remain constant forever. This provides a stark contrast with normative predictions from the representative models studied in Aiyagari eta al (2002), in which the long-run dynamics are driven by exogenous restrictions on the ability of the government to set transfers optimally. This example also indicates that incorporating explicit redistributive motive may lead to substantially different implications about the optimal response to the aggregate shocks.

A third set of results concerns higher frequency implications, in particular, the nature of optimal government policy in booms and recessions. What we have to say about this comes from a version of our model calibrated to capture what we take to be key stylized facts that during recent recessions (1) the left tail of the cross-section distribution of labor income falls by more than right tail and (2) interest rates fall. When we calibrate to fit those targets, we find that in recessions accompanied by higher inequality, it is optimal to increase taxes, transfers, and issue government debt. These effects differ substantially both qualitatively and quantitatively from the counterparts in either a representative agent model or our model were a recession is modeled to be a pure TFP shock that leaves the distribution of skills unchanged. Secondly recessions with low interest rates, thought not critical for short run, affects the transient dynamics and long run properties of optimal policy. Starting from a net asset position (government vis-a-vis the high skilled agent) that implies a 60% debt - to - gdp ratio, the benchmark economy with procylical interest rates shows no discernable trend in net asset positions over a time as long as 5000 years. On the other hand, variants with countercyclical interest rates (that effectively ignore discount factor shocks) repay this debt and starts accumulating assets, albeit it takes about 2500 years to do so.

Here is how we construct these results. After section 2 describes preferences, technologies, endowments, and information flows, section 3 sets the stage for our subsequent results. Theorem 1 exploits the fact that the set of feasible allocations is invariant with respect to changes in transfers and *gross* asset positions to show the extension of the Ricardian equivalence and importance of net assets. We first begin with analyzing the quasi-linear case in section 4. Sections 5 characterizes the the planner maximization problem and obtains a recursive representation. Section 6 characterizes the long run allocations, while Section 7 provides numerical simulations of the optimal policy to business cycle-like shocks.

Draft25Graphs/affine_taxes.pdf

Figure 1: The U.S. tax-transfer system is poorly approximated by a linear function, better by an afine function.

1.1 Relationships to literatures

One a fundamental level our paper is related to both Barro (1974), who showed Ricardian equivalence in a representative agent economy with lump sum taxes, and Barro (1979), who studied the optimal taxation in the same economy when lump sum taxes are ruled out. With incomplete markets and heterogeneity both of the forces uncovered by Barro play role, although the desire for redistribution leads to a richer prescriptions for the optimal policy.

A large literature on Ramsey problems exogenously rules out transfers in the context of representative agent, general equilibrium models. Lucas Jr. and Stokey (1983), Chari et al. (1994), and AMSS are leading examples of this approach. In contrast to those papers, our Ramsey planner cares about redistribution among agents with different skills and wealths. Other than prohibiting them from depending explicitly on agent's personal identity, we leave transfers unrestricted and have the Ramsey planner set them optimally. Nevertheless, we find that some of the general principles that emerge from that 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 substantially changes qualitative and quantitative insights about the optimal policy in important respects.

[1]

[1] There is also a more recent strand of literature that focuses on the optimal policy in settings with heterogeneous

Several recent papers impute distributive concerns to a Ramsey planner. Three papers that are perhaps most closely related to ours are Bassetto (1999), Shin (2006), and Werning (2007). Like us, those authors depart from a representative agent assumption by allowing heterogeneity and considering distributional consequences of alternative tax and borrowing policies. The first paper by Bassetto extends the Lucas Jr. and Stokey (1983) environment to include I types of agents who are heterogeneous in their time-invariant labor productivities. There are complete markets and a Ramsey planner has access only to proportional taxes on labor income and history-contingent borrowing and lending. The authors study how the Ramsey planner's vector of Pareto weights influences how he responds to government expenditures and other shocks by adjusting the proportional labor tax and government borrowing to cover expenses while manipulating prices in ways that redistribute wealth between 'rentiers' (who have low productivities) 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 economy to have two risk-averse households who face idiosyncratic income risk. When idiosyncratic income risk is big enough relative to aggregate government expenditure risk, the Ramsey planner chooses to issue debt in order to help households engage in precautionary saving, thereby overturning the AMSS result that in their quasi-linear case a Ramsey planner eventually sets taxes to zero and lives off its earnings from assets forevermore. Shin emphasizes that the government does this at the cost of imposing tax distortions. While being confined to 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 results about net versus gross asset positions, including that government assets can be set to zero in all periods. Because he allows unrestricted taxation of initial assets, the initial distribution of assets plays no role. Theorem 1 and its corollaries substantially generalize Werning's results by showing that all allocations of assets among agents and the government that imply the same optimal net asset position lead to the same optimal allocation, a conclusion that holds for market structures beyond complete markets. 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 not present when markets are complete.[2] [3]

Finally, our numerical analysis in Section 7 is related to a recent paper by McKay and Reis (2013).

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.

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

[3]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.

While the focus of the two papers is very different- McKay and Reis study the effect of calibrated US tax and transfer system on stabilization of output, we focus on the optimal policy responses in a somewhat simpler economy- both papers reach similar conclusions about the importance of transfers and redistribution over business-cycle frequencies.

2 Environment

Exogenous fundamentals of the economy are functions of a shock s_t that follows an irreducible Markov process, where $s_t \in S$ and S is a finite set. We let $s^t = (s_0, ..., s_t)$ denote a history of shocks.

There is a mass π_i of a type $i \in I$ agent, with $\sum_{i=1}^{I} \pi_i = 1$. Types differ by their skills. Preferences of an agent of type i over stochastic processes for consumption $\{c_{i,t}\}_t$ and labor supply $\{l_{i,t}\}_t$ are ordered by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \left[\Pi_{j=0}^{t} \beta(s_j)\right] U^i \left(c_i(s^t), l_i(s^t)\right) \tag{1}$$

where $\mathbb{E}_t$ is a mathematical expectations operator conditioned on time t information and $\beta(s_t) \in (0,1)$ is a state-dependent discount factor[4]. We assume that $l_i \in \left[0, \bar{l}_i\right]$ for some $\bar{l}_i < \infty$. Results in section 3 require no additional assumptions on U^i like differentiability or convexity[5] , but results in later sections do.

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

$$\sum_{i=1}^{I} \pi_i c_i(s^t) + g\left(s_t\right) = \sum_{i=1}^{I} \pi_i \theta_i\left(s_t\right) l_i(s^t), \tag{2}$$

where $g\left(s_t\right)$ denotes exogenous government expenditures in state s_t. We allow s_t to affect $\beta(s_t)$, government expenditures $g(s_t)$, and the type-specific productivities $\theta_i(s_t)$.

To save on notation, mostly we use z_t to denote a random variable with a time t conditional distribution that is a function of the history s^t. Occasionally, we use the more explicit notion $z\left(s^t\right)$ to denote a realization at a particular history s^t.

A Ramsey planner's preferences over a vector of stochastic processes for consumption and work are ordered by

$$\mathbb{E}_0 \sum_{i=1}^{I} \pi_i \alpha_i \sum_{t=0}^{\infty} \bar{\beta}_t U_t^i \left(c_{i,t}, l_{i,t}\right) \tag{3}$$

where the Pareto weights satisfy $\alpha_i \geq 0$, $\sum_{i=1}^{I} \alpha_i = 1$ and $\bar{\beta}_t = \left[\Pi_{j=0}^{t} \beta_j\right]$

In most of this paper, we study an optimal government policy when agents can trade only a one-period risk-free bond. We assume that the government imposes an affine tax

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

[4] We allow the discount factor to depend on the Markov state s to generate flexible comovement patterns between real interest rates and fundamentals.

[5] Consequently our setup allows both extensive and intensive responses of labor.

We do not restrict the sign of T_t at any t or s^t. If for some type i, $\theta_{i,t}=0$, $b_{i,-1}=0$ and U^i is defined only on $\mathcal{R}^2_+$, his budget constraint will imply that the all feasible allocations for the planner have nonnegative present value of transfers, since transfers are the sole source of a type i agent's wealth and consumption.

While results in sections 5, 6, and 7 depend on these assumptions about an affine tax system and incomplete markets, key results of section 3 apply under more general tax functions and market structures.

Under an affine tax system, agent i's budget constraint at t is

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

where $b_{i,t}$ denotes asset holdings of a type i agent at time $t\geq 0$, R_{t-1} is a gross risk-free one-period interest rate from $t-1$ to t for $t\geq 1$, and $R_{-1}\equiv 1$. For $t\geq 0$, R_t is measurable with respect to s^t. To rule out Ponzi schemes, we assume that $b_{i,t}$ must be bounded from below. Except in subsection 3.1, we impose no further constraints on agents' borrowing and lending. Subsection 3.1 briefly studies economies with arbitrary borrowing constraints.

The government budget constraint is

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

where B_t denotes the government's assets at time t, which we assume are bounded from below. Our assumptions about preferences imply that the government can collect only finite revenues in each period, so this restriction rules out government-run Ponzi schemes.

We assume that private agents and the government start with assets $\{b_{i,-1}\}_{i=1}^{I}$ and B_{-1}, respectively. Asset holdings satisfy the market clearing condition

$$\sum_{i=1}^{I}\pi_i b_{i,t}+B_t=0 \text{ for all } t\geq -1. \tag{6}$$

Since B_t and all $b_{i,t}$ are bounded from below, equation (6) implies that they are also bounded from above.

Components of competitive equilibria are described below

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* price system *is an interest rate sequence* $\{R_t\}_t$. *A* tax policy *is a sequence* $\{\tau_t,T_t\}_t$.

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* $\{c_{i,t},l_{i,t},b_{i,t}\}_{i,t}$ *maximize (1) subject to (4) and* $\{b_{i,t}\}_{i,t}$ *is bounded; and constraints (2), (5) and (6) are satisfied.*

Lastly we define optimal competitive equilibria.

Definition 3 *Given* $(\{b_{i,-1}\}_i,B_{-1})$, *an optimal competitive equilibrium with affine taxes is a tax policy* $\{\tau_t^*,T_t^*\}_t$, *an allocation* $\left\{c_{i,t}^*,l_{i,t}^*\right\}$, *an asset profile* $\left\{\left\{b_{i,t}^*\right\}_i,B_t^*\right\}_t$, *and a price system* $\{R_t^*\}_t$ *such* *that (i) given $\left(\{b_{i,-1}\}_i, B_{-1}\right)$, the tax policy, the price system, and the allocation constitute a competitive equilibrium; and (ii) there is no other tax policy $\{\tau_t, T_t\}_t$ such that a competitive equilibrium given $\left(\{b_{i,-1}\}_i, B_{-1}\right)$ and $\{\tau_t, T_t\}_t$ has a strictly higher value of (3).*

We call $\{\tau_t^*, T_t^*\}_t$ an *optimal tax policy*, $\{c_{i,t}^*, l_{i,t}^*\}_{i,t}$ an *optimal allocation*, and $\left\{\left\{b_{i,t}^*\right\}_i, B_t^*\right\}_t$ an *optimal asset profile.*

3 Relevant and Irrelevant Aspects of the Distribution of Government Debt

This section sets forth a result that underlies much of the analysis in this paper, namely, that the level of government debt is not a state variable for our economy. The reason is that there is an equivalence class of tax policies and asset profiles that support the same competitive equilibrium allocation. A competitive equilibrium allocation pins down only net asset positions. The assertions in this section apply to all competitive equilibria, not just the optimal ones that will be our focus in subsequent sections.

Theorem 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. For any bounded sequences $\left\{\hat{b}_{i,t}\right\}_{i,t\geq -1}$ that satisfy*

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

there exist sequences $\left\{\hat{T}_t\right\}_t$ and $\left\{\hat{B}_t\right\}_{t\geq -1}$ that satisfy (6) such that $\left\{\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_i, \hat{B}_t, R_t\right\}_t$ and $\left\{\tau_t, \hat{T}_t\right\}_t$ constitute a competitive equilibrium given $\left(\left\{\hat{b}_{i,-1}\right\}_i, \hat{B}_{-1}\right)$.

Proof. Let

$$\hat{T}_t = T_t + \left(\hat{b}_{1,t} - b_{1,t}\right) - R_{t-1}\left(\hat{b}_{1,t-1} - b_{1,t-1}\right) \text{ for all } t \geq 0. \tag{7}$$

Given a tax policy $\left\{\tau_t, \hat{T}_t\right\}_t$, the allocation $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_t$ is a feasible choice for consumer i since it satisfies

$$\begin{aligned} c_{i,t} &= (1-\tau_t)\,\theta_{i,t}l_{i,t} + R_{t-1}b_{i,t-1} - b_{i,t} + T_t. \\ &= (1-\tau_t)\,\theta_{i,t}l_{i,t} + R_{t-1}\left(b_{i,t-1} - b_{1,t-1}\right) - \left(b_{i,t} - b_{1,t}\right) + T_t + R_{t-1}b_{1,t-1} - b_{1,t} \\ &= (1-\tau_t)\,\theta_{i,t}l_{i,t} + R_{t-1}\left(\hat{b}_{i,t-1} - \hat{b}_{1,t-1}\right) - \left(\hat{b}_{i,t} - \hat{b}_{1,t}\right) + T_t + R_{t-1}b_{1,t-1} - b_{1,t} \\ &= (1-\tau_t)\,\theta_{i,t}l_{i,t} + R_{t-1}\hat{b}_{i,t-1} - \hat{b}_{i,t} + \hat{T}_t. \end{aligned}$$

Suppose that $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_t$ is not the optimal choice for consumer i, in the sense that there exists some other sequence $\left\{\hat{c}_{i,t}, \hat{l}_{i,t}, \hat{b}_{i,t}\right\}_t$ that gives strictly higher utility. Then the choice $\left\{\hat{c}_{i,t}, \hat{l}_{i,t}, b_{i,t}\right\}_t$ is feasible given the tax rates $\{\tau_t, T_t\}_t$, which contradicts the assumption that $\{c_{i,t}, l_{i,t}, b_{i,t}\}_t$ is the optimal choice for the consumer given taxes $\{\tau_t, T_t\}_t$. The new allocation satisfies all other constraints and therefore is an equilibrium. ■

An immediate corollary is that it is not total government debt but rather who owns it that affects equilibrium allocations.

Corollary 1 *For any pair B'_{-1}, B''_{-1}, there are asset profiles $\left\{b'_{i,-1}\right\}_i$ and $\left\{b''_{i,-1}\right\}_i$ such that equilibrium allocations starting from $\left(\left\{b'_{i,-1}\right\}_i, B'_{-1}\right)$ and from $\left(\left\{b''_{i,-1}\right\}_i, B''_{-1}\right)$ are the same. These asset profiles satisfy*

$$b'_{i,-1} - b'_{1,-1} = b''_{i,-1} - b''_{1,-1} \text{ for all } i.$$

This result is closely related to Ricardian Equivalence in Barro (1974). There are however some important distinctions. In Barro's representative agent model lump sum taxes are not distortionary. In our economy, since the planner does not have person-specific taxes, a lump sum transfer introduces distortions in inequality and as we will see in following sections this force has a significant effect on optimal policy. Despite this, Ricardian equivalence continues to hold.[6] Theorem 1 shows that many different transfer sequences $\{T_t\}_t$ and asset profiles $\{b_{i,t}, B_t\}_{i,t}$ support the same equilibrium allocation. For example, one can set government assets $B_{i,t} = 0$ without loss of generality. Alternatively, we can normalize assets $b_{i,t}$ of any type i.

Theorem 1 continues to hold in more general environments. For example, we could allow agents to trade all Arrow securities and still show that equilibrium allocations depend only on agents' net assets positions. Similarly, our results hold in economies with capital.

3.1 Extension : Borrowing constraints

Representative agent models rule out Ricardian equivalence either by assuming distorting taxes or by imposing ad hoc borrowing constraints. By way of contrast, we have verified that Ricardian equivalence holds in our economy even though there are distorting taxes. Imposing ad-hoc borrowing limits also leaves Ricardian equivalence intact in our economy.[7] In economies with exogenous borrowing constraints, agents' maximization problems include the additional constraints

$$b_{i,t} \geq \underline{b}_i \tag{8}$$

for some exogenously given $\{\underline{b}_i\}_i$.

Definition 4 *For given $\left(\{b_{i,-1}, \underline{b}_i\}_i, B_{-1}\right)$ and $\{\tau_t, T_t\}_t$, a competitive equilibrium with affine taxes and exogenous borrowing constraints is a sequence $\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, B_t, R_t\}_t$ such that $\{c_{i,t}, l_{i,t}, b_{i,t}\}_{i,t}$ maximizes (1) subject to (4) and (8), $\{b_{i,t}\}_{i,t}$ are bounded, and constraints (2), (5) and (6) are satisfied.*

We can define an *optimal* competitive equilibrium with exogenous borrowing constraints by extending Definition 3.

[6] Wallace (1981)'s Modigliani-Miller theorem for a class of government open market operations has a similar flavor. Sargent (1987) describes the structure of a set of related Modigliani-Miller theorems for government finance.

[7] Bryant and Wallace (1984) describe how a government can use borrowing constraints as part of a welfare-improving policy to finance exogenous government expenditures.Sargent and Smith (1987) describe Modigliani-Miller theorems for government finance in a collection of economies in which borrowing constraints on classes of agents produce the kind of rate of return discrepancies that Bryant and Wallace manipulate.

The introduction of the ad-hoc debt limits leaves unaltered the conclusions of Corollary 1 and the role of the initial distribution of assets across agents. The next proposition asserts that ad-hoc borrowing limits do not limit a government's ability to respond to aggregate shocks.[8]

Proposition 1 *Given an initial asset distribution* $\left(\{b_{i,-1}\}_i, B_{-1}\right)$, *let* $\{c_{i,t}, l_{i,t}\}_{i,t}$ *and* $\{R_t\}_t$ *be a competitive equilibrium allocation and interest rate sequence in an economy without exogenous borrowing constraints. Then for any exogenous constraints* $\{\underline{b}_i\}_i$, *there is a government tax policy* $\{\tau_t, T_t\}_t$ *such that* $\{c_{i,t}, l_{i,t}\}_{i,t}$ *is a competitive equilibrium allocation in an economy with exogenous borrowing constraints* $\left(\{b_{i,-1}, \underline{b}_i\}_i, B_{-1}\right)$ *and* $\{\tau_t, T_t\}_t$.

Proof. Let $\{c_{i,t}, l_{i,t}, b_{i,t}\}_{i,t}$ be a competitive equilibrium allocation without exogenous borrowing constraints. Let $\Delta_t \equiv \max_i \{\underline{b}_i - b_{i,t}\}$. Define $\hat{b}_{i,t} \equiv b_{i,t} + \Delta_t$ for all $t \geq 0$ and $\hat{b}_{i,-1} = b_{-1}$. By Theorem 1, $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_{i,t}$ is also a competitive equilibrium allocation without exogenous borrowing constraints. Moreover, by construction $\hat{b}_{i,t} - \underline{b}_i = b_{i,t} + \Delta_t - \underline{b}_i \geq 0$. Therefore, $\hat{b}_{i,t}$ satisfies (8). Since agents' budget sets are smaller in the economy with exogenous borrowing constraints, and $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_{i,t}$ are feasible at interest rate process $\{R_t\}_t$, then $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_{i,t}$ is also an optimal choice for agents in the economy with exogenous borrowing constraints $\{\underline{b}_i\}_i$. Since all market clearing conditions are satisfied, $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_{i,t}$ is a competitive equilibrium allocation and asset profile. ■

To explore the intuition underlying Proposition 1, suppose to the contrary that the exogenous borrowing constraints restricted a government's ability to achieve a desired allocation. That means that the government would want to increase its borrowing and to repay agents later, which the borrowing constraints prevent. But the government can just reduce transfers today and increase them tomorrow. That would achieve the desired allocation without violating the exogenous borrowing constraints.

Welfare can be strictly higher in an economy with exogenous borrowing constraints because a government might want to push some agents against their borrowing limits. When some agents' borrowing constraints bind, their shadow interest rates differ from the common interest rate that unconstrained agents face. When the government rearranges tax policies to affect the interest rate, it affects constrained and unconstrained agents differently. By facilitating redistribution, this can improve welfare. In appendix 9.1, we construct an example without any shocks in which the government can achieve higher welfare by using borrowing constraints to improve its ability to redistribute.

3.2 Ricardian irrelevance and optimal equilibria

Our statements about Ricardian irrelevance apply to all competitive equilibrium allocations, not just the optimal ones that are the main focus of this paper. To appreciate how these Ricardian irrelevance results affect optimal equilibria, suppose that we increase an initial level of government debt from 0 to some arbitrary level $B'_{-1} > 0$. If the government were to hold transfers $\{T_t\}_t$ fixed, it would have to increase tax rates $\{\tau_t\}_t$ enough to collect a present value of revenues sufficient to repay B'_{-1}. Since deadweight

[8] See **?**Yared (2013) who shows a closely related result.

losses are convex in τ, higher levels of debt financed with bigger distorting taxes $\{\tau_t\}$ impose larger distortions on the economy, thereby degrading the equilibrium allocation. But this would not happen if the government were instead to adjust transfers in response to a higher initial debt. To determine optimal transfers, we need to know who owns the initial government debt B'_{-1}. For example, suppose that agents hold equal amounts of it. Then each unit of debt repayment achieves the same redistribution as one unit of transfers. If the original tax policy at $B'_{-1} = 0$ were optimal, then the best policy for a government with initial debt $B'_{-1} > 0$ would be to reduce the present value transfers by exactly the amount of the increase in per capita debt, because then distorting taxes $\{\tau_t\}$ and the allocation would both remain unchanged.[9]

But the situation would be different if holdings of government debt were not equal across agents. For example, suppose that richer people owned disproportionately more government debt than poorer people. That would mean that inequality is effectively initially higher in an economy with higher initial government debt. As a result, a government with Pareto weights $\{\alpha_i\}$ that favor equality would want to increase both distorting tax rates $\{\tau_t\}$ and transfers $\{T_t\}$ to offset the increase in inequality associated with the increase in government debt. The conclusion would be the opposite if government debt were to be owned mostly by poorer households.

This logic shows how important it is to know the distribution of government debt across people. Government debt that is widely distributed across households (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).[10]

4 Quasi-linear preferences

Before we use section 5 for general characterization of our problem, in this section we study a special case of our economy that allows us to get a long way analytically and to identify some forces that drive outcomes. In particular we assume that the only source of aggregate shocks is the fluctuation in g_t and preferences are given by $U^i(c,l) = c - h_i(l)$ where h_i is an increasing differentiable function with $h'_i(0) = 0$ and $h'_i(\bar{l}_i) = \infty$. These quasi-linear preferences have been extensively studied in the context of representative agent economies (see, e.g., AMSS, Farhi (2010), Battaglini and Coate (2007, 2008), **?**, Faraglia et al. (2012)). We pursue two goals in this section. First, this economy provides a stark contrast of the optimal policy in our economy when transfers are chosen optimally with representative agent models in which the choice of transfers is exogenously restricted. Second, this set up switches off two channels which are present more generally, namely, that the marginal utilities of agents are differently affected by changes in transfers and that the interest rate in general is not constant. These two forces will play an important role in determining the long run allocations in Section 6

[9] This example illustrates principles proclaimed by Simon Newcomb (1865, p. 85) in the quotation with which we began this paper.

[10] It is straightforward 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.

To simplify notation, we now assume that the initial debt is $\left\{\beta^{-1} b_{i,-1}\right\}_i$.

Proposition 2 *Suppose that preferences are quasi-linear and the only aggregate shocks are g_t. Then the optimal tax rate τ_t^* satisfies $\tau_t^* = \tau^*$. An optimum asset profile $\left\{b_{i,t}^*, B_t^*\right\}_{i,t}$ can be chosen to satisfy $b_{i,t}^* = b_{i,-1}$ for all $i,\ t \geq 0$ and $B_t^* = B_{-1}$ for all $t \geq 0$.*[11]

Proof. When preferences are quasilinear, the interest rate $R_t = \beta^{-1}$ for all t and $(1-\tau_t)\,\theta_i = h_i'(l_{i,t})$ for all t. For our purposes, it is more convenient to express the labor supply component of the allocation as a function of $(1-\tau)$ and optimize with respect to τ rather than $\{l_i\}_i$. We invert $h_i'(\cdot)$ to express labor supply l_i as a function of $(1-\tau)$. Call this function $H_i(1-\tau)$. Use the budget constraint (4) to obtain

$$c_{i,t} + b_{i,t} - (1-\tau_t)\, H_i(1-\tau_t) = T_t + \beta^{-1} b_{i,t-1}. \tag{9}$$

The optimal allocation solves

$$\max_{\{c_{i,t}, b_{i,t}, \tau_t, T_t\}_{i,t}} \mathbb{E}_0 \sum_{t=0}^{\infty} \sum_{i=1}^{I} \alpha_i \pi_i \beta^t \left[c_{i,t} - h_i\left(H_i(1-\tau_t)\right)\right] \tag{10}$$

subject to $\{b_{i,t}\}_{i,t}$ being bounded, (9), and

$$\sum_{i=1}^{I} \pi_i c_{i,t} + g_t = \sum_{i=1}^{I} \pi_i H_i(1-\tau_t).$$

Note that since $\{b_{i,t}\}_{i,t}$ is bounded,

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\beta^{-1} b_{i,t-1} - b_{i,t}\right] = \beta^{-1} b_{i,-1} + \lim_{\mathcal{T}\to\infty} \mathbb{E}_0 \left(\sum_{t=0}^{\mathcal{T}} \beta^t \left[b_{i,t} - b_{i,t}\right] - \beta^{\mathcal{T}+1} b_{i,\mathcal{T}+1}\right) = \beta^{-1} b_{i,-1}.$$

Use (9) to eliminate $c_{i,t}$ and then use the preceding expression to get

$$\max_{\{b_{i,t}, \tau_t, T_t\}_{i,t}} \mathbb{E}_0 \sum_{t=0}^{\infty} \sum_{i=1}^{I} \alpha_i \pi_i \beta^t \left[T_t + (1-\tau_t)\, H_i(1-\tau_t) - h_i\left(H_i(1-\tau_t)\right)\right] + \beta^{-1} \sum_{i=1}^{I} b_{i,-1}. \tag{11}$$

subject to

$$\sum_{i=1}^{I} \pi_i \left[T_t + \beta^{-1} b_{i,t-1} - b_{i,t} + (1-\tau_t)\, H_i(1-\tau_t)\right] + g_t = \sum_{i=1}^{I} \pi_i H_i(1-\tau_t). \tag{12}$$

Let $\beta^t \lambda_t$ be the Lagrange multiplier on the time t feasibility constraint (2). The first-order condition with respect to T_t implies that $\lambda_t = \sum_{i=1}^{I} \alpha_i \pi_i$ is constant and independent of t. Therefore, optimal taxes $\tau_t = \tau^*$ are also constant and independent of t. Using τ^*, equation (12) pins down $\sum_{i=1}^{I} \pi_i \left[T_t + \beta^{-1} b_{i,t-1} - b_{i,t}\right]$. Without loss of generality we can set $b_{i,t}^* = b_{i,-1}$ and T_t^* to satisfy (12).
■

In the optimal equilibria for the quasi-linear economy described in Proposition 2, fluctuations in lump-sum taxes and transfers "do all the work". In period 0, the government chooses an optimal present value

[11] We thank Guy Laroque for suggesting the idea for this proof.

of transfers and a constant tax rate that pays for it. Tax rates and transfers depend on the Pareto weights $\{\alpha_i\}$: higher Pareto weights on low skilled agents imply higher transfers and tax rates. In response to a shock g_t, the government adjusts transfers in period t by the amount of the shock. Since all agents are risk-neutral, welfare is unaffected by fluctuations in transfers. This allows the government to perfectly smooth distorting tax rates.

Comparison with representative agent economies

The Lucas Jr. and Stokey (1983) and AMSS representative agent models impose $T_t \geq 0$. An informal justification behind doing so is the desire of the government to not hurt poor agents, who might be unable to afford lump-sump taxes. This constraint always binds in the Lucas and Stokey model and that binds in the AMSS model until the government has acquired enough assets to finance all future expenditures from earnings on those assets. In those representative agent models, the government would like to impose lump-sum *taxes*, not transfers. We explicitly model redistributive concerns by imputing Pareto weights to heterogeneous agents and obtained very different dynamics as shown in Proposition 2. In this section we argue that the differences in the dynamics come from the presence of this arbitrary restrictions on transfers and not explicit or implicit redistributory motives.

We impose the following in our maximization problem (10), namely[12],

$$T_t \geq 0 \text{ for all } t. \tag{13}$$

The following proposition states it is optimal for the planner to set policy in such a way that constraint (13) becomes slack overtime[13].

Proposition 3 *Assume that $I \geq 1$ and g_t takes more than one value. Let $\beta^t \chi_t$ be the Lagrange multiplier on constraint (13) in a version of maximization problem (10) that is augmented with constraint (13). Then $\chi_t \to 0$ a.s.*

Proof. Our augmented version of the maximization problem (10) can be expressed as maximization problem (11) with an additional constraint (13). The first-order conditions for T_t yield $\sum_{i=1}^{I} \alpha_i \pi_i = \mu_t + \chi_t$, while the first-order condition for $b_{i,t}$ implies $\mu_t = \mathbb{E}_t \mu_{t+1}$. Since $\chi_t \geq 0$, these two conditions imply that χ_t is a nonnegative martingale and therefore χ_t must almost surely converge to a constant. This, in turn, implies that μ_t must almost surely converge to a constant. Then the first-order conditions for τ_t also imply that τ_t must converge a.s. to some τ^*.

Suppose $\chi_t \to \chi^* > 0$. This implies that $T_t \to 0$ and (12) becomes

$$-\beta^{-1} B_{i,t-1} + B_{i,t} + \sum_{i=1}^{I} \pi_i (1-\tau^*) H_i (1-\tau^*) + g_t = \sum_{i=1}^{I} \pi_i H_i (1-\tau^*),$$

[12] This makes AMSS a special case of our economy

[13] It can be shown that if g is not too high and government is sufficiently redistributory (i.e α_i is sufficiently high for low productivity agents), constraints (13) is *always* slack.

where we used (6) to substitute for $\sum_{i=1}^{I} \pi_i b_{i,t}$. If g_t can take more than one value and follows an irreducible Markov process, then for any bound on B_t, we can find a sequence of government expenditures g_t for which this bound will eventually be violated, leading to a contradiction. This implies that $\chi_t \to 0$. ■

Proposition 2 emphasized that in absence of (13) the government uses flutuations in transfers to finance all fluctuations in expenditures. Constraint 13 imposes an asymmetry in using transfers to smooth fluctuations in expenditures. Around zero it is costless to increase transfers by a small amount but infinitely more costly to decrease them by the same amount. This forces the optimal policy to engineers taxes and assets that allows the economy to eventually grow out this constraint.

For quasi-linear preferences, figure 2 compares equilibrium dynamics in a representative agent (AMSS) economy and an economy with two agents, one who is not productive, and Pareto weights chosen to make transfers be positive at all times and along all histories. The sequences of s_t shocks are identical across the two economics. While tax rates converge to zero for the AMSS economy, they are constant for the heterogeneous agent economy.[14].

Figure 2: Taxes in AMSS (solid line) and heterogeneous agent economy (dotted line) with quasi-linear preferences

5 Optimal equilibria with affine taxes

We return to our original problem formulated in section 5. We further assume assume that $U^i : \mathbb{R}_+^2 \to \mathbb{R}$ is concave in $(c, -l)$ and twice continuously differentiable. We let $U_{x,t}^i$ or $U_{xy,t}^i$ denote first and second derivatives of U^i with respect to $x, y \in \{c, l\}$ in period t and assume that $\lim_{x \to \bar{l}_i} U_l^i(c, x) = \infty$, $\lim_{x \to 0} U_l^i(c, x) = 0$ for all c and i.

We focus on interior equilibria. First-order necessary conditions for the consumer's problem are

$$(1 - \tau_t)\, \theta_{i,t} U_{c,t}^i = -U_{l,t}^i \tag{14}$$

and

$$U_{c,t}^i = \beta_t R_t \mathbb{E}_t U_{c,t+1}^i. \tag{15}$$

To help characterize an equilibrium, we use

Proposition 4 *A sequence $\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, R_t, \tau_t, T_t\}_t$ is part of a competitive equilibrium with affine taxes if and only if it satisfies (2), (4), (14), and (15) and $b_{i,t}$ is bounded for all i and t.*

Proof. Necessity is obvious. In the appendix 9.2, we use arguments of Magill and Quinzii (1994) and Constantinides and Duffie (1996) to show that any $\{c_{i,t}, l_{i,t}, b_{i,t}\}_{i,t}$ that satisfies (4), (14), and (15) is

[14] The plots for the AMSS and the heterogeneous agent economies are both for quasi-linear preferences with a Frisch elasticity of labor equals 0.5 and a discount factor $\beta = 0.95$. In the AMSS economy, the agent's initial assets are zero and government expenditure shocks $g(s_t) \in \{.1, .3\}$ are generated using an IID process with equally likely outcomes. For the heterogeneous agent economy, we set $\alpha_2 = .54$ so that the initial labor taxes are similar to those for the AMSS economy

a solution to consumer i's problem. Equilibrium $\{B_t\}$ is determined by (6) and constraint (5) is then implied by Walras' Law ■

To find an optimal equilibrium, by Proposition 4 we can choose $\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, R_t, \tau_t, T_t\}_t$ to maximize (3) subject to (2), (4),(14), and (15). We apply a first-order approach and follow steps similar to ones taken by Lucas Jr. and Stokey (1983) and AMSS. Substituting consumers' first-order conditions (14) and (15) into the budget constraints (4) yields implementability constraints

$$c_{i,t} + b_{i,t} = -\frac{U^i_{l,t}}{U^i_{c,t}} l_{i,t} + T_t + \frac{U^i_{c,t-1}}{\beta_{t-1}\mathbb{E}_{t-1}U^i_{c,t}} b_{i,t-1} \text{ for all } i, t. \tag{16}$$

For $I \geq 2$, we can use constraint (16) for $i = 1$ to eliminate T_t from (16) for $i > 1$. Define $\tilde{b}_{i,t} \equiv b_{i,t} - b_{1,t}$ we can represent the implementability constraints as

$$\begin{aligned} & (c_{i,t} - c_{1,t}) + \tilde{b}_{i,t} \\ = \quad & -\frac{U^i_{l,t}}{U^i_{c,t}}{}_{i,t}\; l_{i,t} + \frac{U^1_{l,t}}{U^1_{c,t}} l_{1,t} + \frac{U^i_{c,t-1}}{\beta_{t-1}\mathbb{E}_{t-1}U^i_{c,t}} \tilde{b}_{i,t-1} \text{ for all } i > 1, t. \end{aligned} \tag{17}$$

With this representation of the implementability constraints, the planner's maximization problem depends only on the $I-1$ variables $\tilde{b}_{i,t-1}$. The reduction of the dimensionality from I to $I-1$ is another consequence of theorem 1.

Denote $I^i_t = U^i_{c,t} c_{i,t} + U^i_{l,t}$, by iterating on equation 17 we get

$$\tilde{b}_{t-1} \frac{U^i_{c,t-1}}{\beta_{t-1}\mathbb{E}_{t-1}U^i_{c,t}} = \mathbb{E}_t \sum_{j=0}^{\infty} \left[\prod_{k=0}^{j} \beta_{t+k}\right] \left[\frac{I^i_{t+j} - I^1_{t+j}}{U^i_{c,t}}\right] \quad \forall t \geq 1 \tag{18a}$$

$$\tilde{b}_{i,-1} = \mathbb{E}_{-1} \sum_{j=0}^{\infty} \left[\prod_{k=0}^{j} \beta_{t+k}\right] \left[\frac{I^i_{t+j} - I^1_{t+j}}{U^i_{c,t}}\right] \tag{18b}$$

The left hand side of (18a) is the post interest savings of agent i and it imposes a "measurablity" constraint on the allocations such that the right hand side determined in $t-1$.

We can now write the problem recursively. Let $\boldsymbol{x} = \beta^{-1}\left(U^2_c \tilde{b}_2, ..., U^I_c \tilde{b}_I\right)$, $\boldsymbol{\rho} = \left(U^2_c/U^1_c, ..., U^I_c/U^1_c\right)$, and denote an allocation $a = \{c_i, l_i\}_{i=1}^I$. In the spirit of Kydland and Prescott (1980) and Farhi (2010), we split the problem into a time-0 problem that takes $(\{\tilde{b}_{i,-1}\}_{i=2}^I, s_0)$ as given and a time $t \geq 1$ continuation problem that takes $\boldsymbol{x}, \boldsymbol{\rho}, s_-$ as given. We formulate two Bellman equations and two value functions, one that pertains to $t \geq 1$, another for $t = 0$.

For $t \geq 1$, let $V(\boldsymbol{x}, \boldsymbol{\rho}, s_-)$ be the continuation value to the planner given $\boldsymbol{x}_{t-1} = \boldsymbol{x}, \boldsymbol{\rho}_{t-1} = \boldsymbol{\rho}, s_{t-1} = s_-$. It satisfies the Bellman equation

$$V(\boldsymbol{x}, \boldsymbol{\rho}, s_-) = \max_{a(s), x'(s), \rho'(s)} \sum_s \Pr(s|s_-) \left(\left[\sum_i \pi_i \alpha_i U^i(s)\right] + \beta(s) V(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s), s) \right) \tag{19}$$

subject to

$$U^i_c(s)\left[c_i(s) - c_1(s)\right] + \beta(s) x'_i(s) + \left(U^i_l(s) l_i(s) - U^i_c(s)\frac{U^1_l(s)}{U^1_c(s)} l_1(s)\right) = \frac{x U^i_c(s)}{\mathbb{E}_{s_-} \boldsymbol{U}^i_c} \text{ for all } s, i \geq 2 \tag{20a}$$

$$\frac{\mathbb{E}_{s\_}\boldsymbol{U}_c^i}{\mathbb{E}_{s\_}\boldsymbol{U}_c^1} = \rho_i \text{ for all } i \geq 2 \tag{20b}$$

$$\frac{U_l^i(s)}{\theta_i(s)U_c^i(s)} = \frac{U_l^1(s)}{\theta_1(s)U_c^1(s)} \text{ for all } s, i \geq 2 \tag{20c}$$

$$\sum_i \pi_i c_i(s) + g(s) = \sum_i \pi_i \theta_i(s) l_i(s) \quad \forall s \tag{20d}$$

$$\rho_i'(s) = \frac{U_c^i(s)}{U_c^1(s)} \text{ for all } s, i \geq 2 \tag{20e}$$

Constraints (20b) and (20e) implies (15). The definition of x_t and constraints (20a) together exhausts equation (17) scaled by U_c^i.

Let $V_0\left(\{\tilde{b}_{i,-1}\}_{i=2}^I, s_0\right)$ be the value to the planner at $t = 0$, where $\tilde{b}_{i,-1}$ denotes initial debt inclusive of accrued interest. It satisfies the Bellman equation

$$V_0\left(\{\tilde{b}_{i,-1}\}_{i=2}^I, s_0\right) = \max_{a_0, x_0, \rho_0} \sum_i \pi_i \alpha_i U^i(c_{i,0}, l_{i,0}) + \beta(s_0) V\left(x_0, \rho_0, s_0\right) \tag{21}$$

subject to

$$U_{c,0}^i\left[c_{i,0} - c_{1,0}\right] + \beta(s_0) x_{i,0} + \left(U_{l,0}^i l_{i,0} - U_{c,0}^i \frac{U_{l,0}^1}{U_{c,0}^1} l_{1,0}\right) = U_{c,0}^i \tilde{b}_{i,-1} \text{ for all } i \geq 2 \tag{22a}$$

$$\frac{U_{l,0}^i}{\theta_{i,0} U_{c,0}^i} = \frac{U_{l,0}^1}{\theta_{1,0} U_c^{1,0}} \text{ for all } i \geq 2 \tag{22b}$$

$$\sum_i \pi_i c_{i,0} + g_0 = \sum_i \pi_i \theta_{i,0} l_{i,0} \tag{22c}$$

$$\rho_{i,0} = \frac{U_{c,0}^i}{U_{c,0}^1} \text{ for all } i \geq 2 \tag{22d}$$

The time 0 problem differs from the time $t \geq 1$ problem since constraint (20b) is absent from the time 0 problem.

6 Ergodic distribution and policies in the long run

In this section, we characterize the properties of the ergodic set to which state variables converge over time. We start with a case when aggregate shocks are iid and can take two values. We show that for this shock structure there generally exists a "steady state" $\left(\boldsymbol{x}^{SS}, \boldsymbol{\rho}^{SS}\right)$ such that if economy ever reaches this state, it stays there. Taxes and transfers in this steady state depend only on the current realization of the shock and the fluctuations of taxes are small for commonly used preferences. We also characterize properties of the steady state, discuss conditions under which economy converges to it and the speed of convergence. Section 6.3 then extends the analysis to more general shocks and shows numerically that while the "steady state" generally does not exist, the properties of the ergodic set are very similar to those in the two shock iid case. Throughout this section we assume that preferences are separable in consumption and labor.

6.1 IID shocks with two values

Let $\Psi\left(s;\boldsymbol{x},\boldsymbol{\rho},s_{-}\right)$ be an optimal law of motion for the state variables for the $t\geq 1$ recursive problem, i.e. $\Psi\left(s;\boldsymbol{x},\boldsymbol{\rho},s_{-}\right)=\left(x'\left(s\right),\rho'\left(s\right)\right)$ that solves (19) given state $\left(\boldsymbol{x},\boldsymbol{\rho},s_{-}\right)$.

Definition 5 *A steady state is* $\left(\boldsymbol{x}^{SS},\boldsymbol{\rho}^{SS}\right)$ *that satisfies* $\left(\boldsymbol{x}^{SS},\boldsymbol{\rho}^{SS}\right)=\Psi\left(s;\boldsymbol{x}^{SS},\boldsymbol{\rho}^{SS},s_{-}\right)$ *for all* s,s_{-}.

Since in this steady state $\rho_i=U_c^i(s)/U_c^1(s)$ does not depend on the realization of shock s, the ratios of marginal utilities of the all agent is constant. The continuation allocation depends only on s_t and not on the history s^{t-1}.

We first begin by noting that the competitive equilibrium implicitly identifies an allocation $\{c_i(s),l_i(s)\}_i$ given a choice for $\{\tau(s),\boldsymbol{\rho}(s)\}$ using equations (20c), (20d) and (20e). Let us denote $U(\tau,\boldsymbol{\rho},s)$ as the value for the planner from the implied allocation using Pareto weights $\{\alpha_i\}_i$,

$$U(\tau,\boldsymbol{\rho},s)=\sum_i\alpha_iU^i(s).$$

Let $I_i(s)=[1-\tau(s)]l_i(s)-c_i(s)$ be the disposable income of agent i in state s, define $Z_i(\tau,\rho,s)$ as the utility adjusted spread in the disposable income relative to Agent 1 :

$$Z_i(\tau,\boldsymbol{\rho},s)=U_c^i(s)\left\{I_1(s)-I_i(s)\right\}.$$

The optimal policy solves the following Bellman equation for $\boldsymbol{x}(s^{t-1})=\boldsymbol{x},\boldsymbol{\rho}(s^{t-1})=\boldsymbol{\rho}$

$$V(\boldsymbol{x},\boldsymbol{\rho})=\max_{\tau(s),\boldsymbol{\rho}'(s),\boldsymbol{x}'(s)}\sum_sP(s)\left[U(\tau(s),\boldsymbol{\rho}'(s),s)+\beta(s)V(\boldsymbol{x}'(s),\boldsymbol{\rho}'(s))\right]\tag{23}$$

subject to the constraints

$$Z_i(\tau(s),\boldsymbol{\rho}'(s),s)+\beta(s)x_i'(s)=\frac{x_iU_c^i(\tau(s),\boldsymbol{\rho}'(s),s)}{\mathbb{E}U_c^i(\tau,\rho)}\text{ for all }s,i\geq 2,\tag{24}$$

$$\sum_sP(s)U_c^1(\tau(s),\boldsymbol{\rho}'(s),s)(\rho_i'(s)-\rho_i)=0\text{ for }i\geq 2.\tag{25}$$

Constraint (25) is obtained by rearranging constraint (20b). It implies that $\rho(s)$ is a risk-adjusted martingale and we use this property later to discuss convergence. We next check if the first order necessary conditions are consistent with stationary policies for some $(\boldsymbol{x},\boldsymbol{\rho})$.[15]

Lemma 1 *Let* $\Pr(s)\mu_i(s)$ *and* λ_i *be the multipliers on constraints (24) and (25). Imposing the restrictions* $x_i'(s)=x_i$ *and* $\rho_i'(s)=\rho_i$ *the steady state solves for* $\{\mu_i,\lambda_i,x_i,\rho_i\}_{i=2}^N$ *and* $\{\tau(s)\}_s$ *using the following equations*

$$Z_i(\tau(s),\boldsymbol{\rho},s)+\beta(s)x_i'=\frac{x_iU_c^1(\tau(s),\boldsymbol{\rho},s)}{\mathbb{E}U_c^1(\tau,\rho)}\textit{ for all }s,i\geq 2,\tag{26a}$$

$$U_\tau(\tau(s),\boldsymbol{\rho},s)-\sum_i\mu_iZ_{i,\tau}(\tau(s),\boldsymbol{\rho},s)=0\textit{ for all }s,\tag{26b}$$

$$U_{\rho_i}(\tau(s),\boldsymbol{\rho},s)-\sum_j\mu_j(s)Z_{j,\rho_i}(\tau(s),\boldsymbol{\rho},s)+\lambda_iU_c^i(\tau(s),\boldsymbol{\rho}'(s),s)-\lambda_i\beta(s)\mathbb{E}U_c^i(\tau,\boldsymbol{\rho})=0.\textit{ for all }s,i\geq 2\tag{26c}$$

[15]Appendix 9.5 discuses the associated second order conditions that ensure these policies are optimal

Since the shock s can take only two values, the system (26) is a square system in $4(N-1)+2$ unknows $\{\mu_i^{SS}, \lambda_i^{SS}, x_i^{SS}, \rho_i^{SS}\}_{i=2}^N$ and $\{\tau^{SS}(s)\}_s$. One can numerically verify that this system has a solution for wide range of primitives. In the next section we formally establish this for a class of simple two agent economies that, while special, illustrates general forces that affect outcomes. The example will help us develop some comparative statics and interpret outcomes from quantitative analysis in section 7.

Lemma 1 also highlights the tradeoffs that the planner faces. Defining $\tilde{\lambda} = -\lambda \mathbb{E} U_c^i(\tau, \boldsymbol{\rho})$ and taking expectations for equation (26c), we get that

$$\mathbb{E} U_{\rho_i}(\tau(s), \boldsymbol{\rho}, s) = \mathbb{E} \sum_j \mu_j(s) Z_{j,\rho_i}(\tau(s), \boldsymbol{\rho}, s) + (1 - \mathbb{E}\beta(s))\tilde{\lambda}_i \tag{27}$$

The multiplier on the implementability constraint for i can be interpreted as the marginal costs of extracting funds from i and $\tilde{\lambda}_i$ is proportional to the multiplier on the constraint $\frac{\mathbb{E}\boldsymbol{U}_c^i}{\mathbb{E}\boldsymbol{U}_c^1} = \rho$. This constraint ensures that at the optimal allocation, agent i has no incentives to participate in the risk free market to change his bond portfolios. The left hand side of (27) captures the cost for the planner if inequality (measured by the ratios of marginal utilities of consumption) deviates from his ideal point, given by α_i/α_1. In the absence of any constraints, the planner would set $\mathbb{E} U_{\rho_i}(\tau(s), \boldsymbol{\rho}, s) = 0$, which implies that $\alpha_i U_c^i = \alpha_1 U_c^1$ for all i. The right hand side of equation (27) captures the costs of approching the planner's ideal point, which come from the costs of raising taxes (the first term on the left hand side) and ability of agents to trade with each other (the second term).

The behavior of the economy in the steady state is similar to the behavior of the complete market economy characterized by Werning (2007). Both taxes and transfers depend only on current relazation of shock s_t. Moreover, the arguments of Werning (2007) can be adapted directly to show that taxes are constant when preferences have a CES form $c^{1-\sigma}/(1-\sigma) - l^{1+\gamma}/(1-\gamma)$ and flucations in taxes is approximately zero when preferences take a balanced growth path form. We return to this point once we discuss convergence properties.

A two agent example

Lemma 5 provides a simple way to verify existence of a steady state for wide range of parameter values by checking that there exists a root for (26). Since the system of equations (26) is non-linear, this existence can generally be verified only numerically. In this section we provide a simple example with risk averse agents in which we can show existence of the root to (26) analytically. The analytical characterization of the steady state will allow us to show two main forces that determine the steady state asset distribution. These forces will also help to understand the long run behavior of the calibrated economy that we study in section 7.

Consider an economy consisting of two types of households with $\theta_{1,t} > \theta_{2,t} = 0$. One period utilities are $\ln c - \frac{1}{2} l^2$. The shock s takes two values, $s \in \{s_L, s_H\}$ with probabilities $\Pr(s|s_{\_})$ that are independent of s_-. We assume that $g(s) = g$ for all s, and $\theta_1(s_H) > \theta_1(s_L)$. We allow the discount factor $\beta(s)$ to depend on s.

Proposition 5 *Suppose that $g < \theta(s)$ for all s.*

1. ***Countercyclical interest rates.*** *If $\beta(s_H) = \beta(s_L)$, then there exists a steady state $\left(x^{SS}, \rho^{SS}\right)$ such that $x^{SS} > 0$, $R^{SS}(s_H) < R^{SS}(s_L)$.*

2. ***Acyclical interest rates.*** *There exists a pair $\{\beta(s_H), \beta(s_L)\}$ such that there exists a steady state with $x^{SS} > 0$ and $R^{SS}(s_H) = R^{SS}(s_L)$.*

3. ***Procyclical interest rates.*** *There exists a pair $\{\beta(s_H), \beta(s_L)\}$ such that there exists a steady state with $x^{SS} < 0$ and $R^{SS}(s_H) > R^{SS}(s_L)$.*

In all cases, taxes $\tau(s) = \tau^{SS}$ are independent of the realized state.

In this two agent case, normalizing assets of the unproductive agent (using theorem 1) we can interpret x as the marginal utility adjusted assets of the government. Besides establishing existence, the proposition identifies the importance of cyclical properties of real interest rates in determining the sign of these assets and thus enables us to compare our results to representative agent economies like AMSS.

Proposition 5 shows two main forces that determine the dynamics of taxes and assets: fluctuations in inequality and fluctuations in the interest rates. Let start with part 2 of proposition 5, which turns off the second force. When interest rates are fixed, the government can adjust two instruments in response to an adverse shock (i.e., a fall in θ_1): it can either increase the tax rate τ or it can decrease transfers T. Both responses are distortionary, but for different reasons. Increasing the tax rate increases distortions because the deadweight loss is convex in the tax rate, as in Barro (1979). This force operates in our economy just as it does in representative agent economies. But in a heterogeneous agent economy like ours, adjusting transfers T is also costly. When agents' asset holdings are identical, a decrease in transfers disproportionately affects a low-skilled agent, so his marginal utility falls by more than does the marginal utility of a high-skilled agent. Consequently, a decrease in transfers increases inequality, a cost not present in representative agent economies.

The government can reduce the costs of inequality distortions by choosing tax rate policies that make the net asset positions of the high skilled agent decrease over time. That makes the two agents' after-tax and after-interest income get closer together, allowing decreases in transfers to have smaller effects on inequality in marginal utilities. If the net asset position of a high skilled agent is sufficiently low, then a change in transfers has no effect on inequality and all distortions from fluctuations in transfers are eliminated.[16]

Turning now to the second force, interest rates generally fluctuate with shocks. Parts 1 and 3 of proposition 5 indicate what drives those fluctuations. Consider again the example of a decrease in productivity of high skilled agent. If the tax rate τ is left unchanged, the government faces a shortfall of revenues. Since g is constant, the government requires extra sources of revenues. But suppose that

[16]This convergence outcome has a similar flavor to "back-loading" results of Ray (2002) and Albanesi and Armenter (2012) that reflect the optimality of structuring policies intertemporally eventually to disarm distortions.

the interest rate increases whenever θ_1 decreases, as happens, for example, when discount factors are constant and θ_1 is the only source of shocks. If the government holds positive assets, its earnings from those assets increase. So holding assets allows higher interest income to offset some of the government's revenue losses from taxes on labor. The situation reverses if interest rates fall at times of increased need for government revenues, as in part 3 of proposition 5 and the steady state allocation features government holding debt.

What matters for our second force is the comovement of the interest rate with fundamentals shocks. States with low average TFP (and therefore a lower base for labor taxes), high g, or a high spread of productivities that threatens to induce higher inequality (and therefore higher transfers and thirst for more government revenues to finance them) are "adverse" from the point of view of current government finance. The government can cope with such adverse states in less distorting ways if finds itself holding positive (negative) assets if interest rates are high (low). [17].

Depending on details of shock processes, these two forces can either reinforce each other (as happens in Part 1 of proposition 5) or work in the opposite direction (as in Part 3 of proposition 5). In the latter case, whether the government ends up with assets or debt in the long run depends on the relative strengths of the two forces.

Besides discount factor shocks, the level of net assets in the steady state depends on other primitives too such as desire for redistribution. An interesting comparative static exercise is shutting off discount factor shocks and increasing α_1, the weight of the high-skilled agent. This implies that the planner taxes less the high skilled agent and redistributes less income to the low skilled agent. Since after-tax income of low skilled agent is lower, the fluctuation in transfers affects this agent more severe that the high skilled. To smooth this flucations the government needs to accumulate more claims on the high skilled, implying a positive relationship between the steady state level of government assets and Pareto weight on high skilled. Figure 3 plots how taxes and assets of the government vary as we change the Pareto weights on high type, α_1.

Figure 3: Stead state assets : $\tilde{b}_2(s) = \frac{\beta x^{SS}}{U_c^2(s)}$ and taxes : τ^{SS} as a function of Agent 1's (high skilled) Pareto weight

6.2 Stability

In this section we return to the general formulation of the problem from section 5 to study convergence to the steady state. We first begin with describing a test for local convergence using a linear approximation of the policy rules at the steady state. Next, we apply this test to show local stability of the steady state for a wide range of parameters. One additional insight that emerges from these examples that converge to the steady state for the commonly used parameter values is very slow.

[17]The results of proposition 5 can be substantially generalized to an economy with no heterogeneity. In a companion paper Bhandari et all (2013), we study a representative agent economy with a more general incomplete market structure and distorting taxes. We show that for a wide range of preferences a steady state exists and it is globally stable and that the sign of long run asset positition of the government is determined by the co-movement of returns on assets and shocks

To study converges we return to the maximization problem (23) and assume that it admits a steady state. As before, let assume that $\Pr(s)\mu_i(s)$ and λ_i be the multipliers on constraints (24) and (25). In Appendix 9.5 we show that the solution in state s^t can be written recursively in terms of $\{\boldsymbol{\mu}(s^{t-1}), \boldsymbol{\rho}(s^{t-1})\}$ and s_t. The solution can be linearized around the steady state using $(\boldsymbol{\mu}, \boldsymbol{\rho})$ as state variables.[18]

Formally, let $\hat{\Psi}_t = \begin{bmatrix} \boldsymbol{\mu}_t - \boldsymbol{\mu}^{SS} \\ \boldsymbol{\rho}_t - \boldsymbol{\rho}^{SS} \end{bmatrix}$ be the deviation from the steady state. From the linear approximation one can ontain $B(s)$ such that

$$\hat{\Psi}(s_{t+1}, \hat{\Psi}_t) = B(s_{t+1})\hat{\Psi}_t. \tag{28}$$

This linearized system has coefficients that are functions of the shock. The next proposition describes a simple numerical test that allows to verify if this linear system converges to zero in probability.

Proposition 6 *If the (real part) of eigenvalues of $\mathbb{E}B(s)$ are less than 1, system (28) convergences to zero in mean. Further for large t, the conditional variance of $\hat{\Psi}$, denoted by $\Sigma_{\Psi,t}$, follows a deterministic process governed by*

$$vec(\Sigma_{\Psi,t}) = \hat{B} vec(\Sigma_{\Psi,t-1})$$

In addition, if the (real part) of eigenvalues of $\hat{B}$ are less than 1, the system converges in probability.

The eigenvalues (in particular the largest or the dominant one) are instructive not only for whether the system is locally stable but also how quickly the steady state is reached. In particular, the half-life of convergence to the steady state is given by $-\frac{\log(0.5)}{(1-\|\iota\|)}$, where $\|\iota\|$ is the absolute value of the dominant eigenvalue. Thus, the closer the dominant eigenvalue is to one, the slower the speed of convergence is.

We used proposition 6 to verify local stability of a wide range of examples. The typical finding is that the steady state is generically stable and the speed of convergence is slow. In figure 4 we plot the comparative statics for the dominant eigenvalue and associated the half-life for a two agent economy with CES preferences with respect to the size of the shock and risk aversion. We set the other parameter to match a Frisch elasticity of 0.5, real interest rate of 2%, marginal tax rates are around 20% and a 90-10 percentile ratio of wage earnings to be 4. In the first exercise, we vary the size of the expenditure shock to generate output falls from 1% to 10% keeping risk aversion (σ) at one. In the bottom panel, we fix the size of shock such that it produces a 3% output fall at $\sigma = 1$ and vary σ from 0.5 to 7. We see that the dominant eigenvalue is everywhere less than one but very close to one, so that the steady state is stable but convergence is slow for resonable values of curvatures and shocks. We comeback to this feature in section 7 where we study low frequency components of government debt.

Figure 4: Eigenvalues of $\hat{B}$ (left panel) and associated half-life (right panel) as a function of Agent 1's (high skilled) Pareto weight

[18] One could in principle look for a solution in state variables $(\boldsymbol{x}(s^{t-1}), \boldsymbol{\rho}(s^{t-1}))$. For $N = 2$ with $\{\theta_i(s)\}$ different across agents, this would give identical policies and a map which is (locally) invertible between $\boldsymbol{x}$ and $\boldsymbol{\mu}$ for a given $\boldsymbol{\rho}$. However in other cases, it turns out there are unique linear policies in $(\boldsymbol{\mu}, \boldsymbol{\rho})$ and not necessarily in $(\boldsymbol{x}, \boldsymbol{\rho})$. This comes from the fact that the set of feasible $(\boldsymbol{x}, \boldsymbol{\rho})$ are restricted at time 0 and may not contain an open set around the steady state values. When we linearize using $(\boldsymbol{\mu}, \boldsymbol{\rho})$ as state variables, the optimal policies for $\boldsymbol{x}(s^t), \boldsymbol{\rho}(s^t)$ converge to their steady state levels for all pertubations in $(\boldsymbol{\mu}, \boldsymbol{\rho})$.

[MG: I propose to cut the rest of this section. I think it is too loose and vague and does not add anything substantive to the points that we already made in this section. I think it is better to put it into BEGS2 where we can actually have tight results about it.]

Our first observation is that optimality conditions of problem (19) imply that the Language multiplier on the implementability constraint (20a) and the ratios of marginal utilities $\rho_{i,t}$ are risk-adjusted martingales.

The FOC for (19) with respect to $x(s)$ gives us

$$\mathbb{E}\frac{U_c^i(s)}{\mathbb{E}\boldsymbol{U}_c^i}\mu_i(s) = \mu_i \tag{29a}$$

$$\mathbb{E}\mu_i(s) = \mu_i - cov\left[\mu_i(s), \frac{U_c^i(s)}{\mathbb{E}U_c^i(s)}\right] \tag{29b}$$

Similarly, rearranging the bond pricing equation (20b) implies

$$\mathbb{E}\frac{U_c^1(s)}{\mathbb{E}\boldsymbol{U}_c^1}\rho_i(s) = \rho_i \tag{30a}$$

$$\mathbb{E}\rho_i(s) = \rho_i - cov\left[\rho_i(s), \frac{U_c^1(s)}{\mathbb{E}U_c^1(s)}\right] \tag{30b}$$

The first margingale shows up in the representative agent incomplete market models and captures that idea that the planner want to smooth fluctuations in distortions from taxes over time. The second martingale is new. It shows that fluctuations in inequality also follow a risk-adjusted martinage process.

The sign of the covariance terms in equations (29b) and (30b) imparts the drift to the variables towards the steady state. For instance, consider the economy with TFP shocks and a government that starts off with asset distribution skewed towards the high productivity agent (w.l.o.g say $\theta_i(s) < \theta_1(s)$) such that it implies $\mu_{i,t} > \mu_i^{SS}$ and $\rho_{i,t} > \rho_i^{SS}$. With two states, determining the sign of the covariance terms is equivalent to ordering $\mu_i(s)$ and $\rho_i(s)$ relative to $c_i(s)$. Suppose consumption for each agent is procyclical or $c_i(s_l) < c_i(s_h)$. In order to converge to the steady state, it should be the case that both the covariances are positive or $\mu_i(s)$ and $\rho_i(s)$ are countercyclical for such initial conditions. The envelope theorem implies that $\mu_i(s) = V_{x_i}(\boldsymbol{x}(s), \boldsymbol{\rho}(s))$ which can be interpreted as the value of an extra unit of asset for agent i in state s. With countercyclical interest rates (and low initial assets), a dollar is more valuable to agent i in low TFP states. Thus we can expect $\mu_i(s)$ to be countercyclical. Next for $\rho_i(s)$ to be countercyclical as well, we need $\frac{U_c^1(s_l)-U_c^1(s_h)}{U_c^1(s_h)} < \frac{U_c^i(s_l)-U_c^i(s_h)}{U_c^1 i(s_h)}$. This means for low levels of consumption, the relative flucations in marginal utilities of agent i should be larger than that of the agent 1 who has a high present value of earnings. Both these conditions are intuitive and verified in our numerical examples in section 7. Further the linear policies described above also preserve these orderings and as long as the steady state is unique, one can expect these to hold even for the region outside the neighborhood of steady state where these policies are approximated.

6.3 More general shocks

The results on existence and convergence of steady state relied on a special binary-IID restriction. When there are more than two possible values for the shocks or when shocks are persistent, the time-invariant steady state will no longer exist. Mathematically, this occurs because one asset and one risk-free rate of return cannot span all possible needs for government revenues. With richer shock structures, there exists an attraction region in the (x, ρ) space to which the dynamic system converges. Although (x, ρ) are no longer constant in such region, their fluctuations tend to be markedly reduced relative to the transients that occur away from that region, and general properties of x and ρ are the same as those described in Proposition 5. Figure 5 shows long sample paths for economies hit by more general TFP shocks. The top panel has IID shocks with 2 (bold) and 3 (dotted) possible values and the bottom panel has persistent shocks with 2 (bold) and 3 (dotted)possible values.

Figure 5: The figure depicts sample paths of marginal utility adjusted debt of the government i.e $-x_t$. The top panel has IID shocks with 2 (bold) and 3 (dotted) possible values and the bottom panel has persistent shocks with 2 (bold) and 3 (dotted)possible values

7 Optimal policy in booms and recessions

In section 6 we used steady states to characterize the long run behavior of optimal allocations and forces that guide the asymptotic level of net assets. In this section, we use a calibrated version of the economy to a) revisit the magnitude of these forces and b) study optimal policy responses over business cycle frequencies when the economy is possibly far away from the steady state. We choose shocks to match stylized facts about recent recessions in US.

In particular we consider an economy with two types of agents of equal measures with preferences[19]

$$U\left(c,l\right)=\psi\ln c+\left(1-\psi\right)\ln\left(1-l\right).$$

The shock s takes two values, s_H and s_L, and follows a persistent process. We allow β, θ_i and g to be functions of s. We first pick $\bar{\theta}_i, \bar{g}$ and $\bar{\beta}$ for a deterministic economy without shocks and calibrate (ψ, α) to some low frequency data moments. Then to match some business cycle moments we pick shocks according to

$$\begin{aligned}\theta_i(s) &= \bar{\theta}_i[1+\hat{\theta}_i(s)],\\ \beta\left(s\right) &= \bar{\beta}\left[1+\hat{\beta}(s)\right],\\ g\left(s\right) &= \bar{g}\left[1+\hat{g}\left(s\right)\right],\end{aligned} \tag{31}$$

[19]We restrict our attention to the economy with two agents for computational tractability. We want to understand both short-run and long-run responses to shocks. For some of our computations, it is important to allow our dynamic systems to travel over a large subset of state space, including regions encountered infrequently in the invariant distribution. With more agents, it seems possible to apply other methods, for example those of Judd et al. (2011), to study dynamics of our economy within its invariant distribution. We hope to pursue such extensions in future work.

where $\hat{\theta}_i(s) \in \{-e_{i,\theta}, e_{i,\theta}\}$, $\hat{\beta}(s) \in \{-e_\beta, e_\beta\}$ and $\hat{g}(s) \in \{-e_g, e_g\}$. Throughout our experiments, we normalize $b_{2,t} = 0$ for all $t \geq -1$. From market clearing, $B_t = -b_{1,t}$. We refer to B_t as government debt (when negative) and assets (when positive).

7.1 Calibration

We calibrate the model in two steps. We first chose baseline parameters that govern preferences and technology so that an optimal equilibrium for the static[20] version of the economy matches some sample moments in post war US data. In the second step, we adjusted other parameters to make the amplitudes of fluctuations equal to average peak-trough spreads observed in the three most recent recessions (1991-92, 2001-02 and 2008-10).

We first discuss calibration of $\left(\psi, \alpha, \bar{\theta}_i, \bar{g}, \bar{\beta}\right)$. Although these parameters jointly determine the relevant moments, it is helpful to explain which moment in the data mainly influences each parameter. We normalize $\bar{\theta}_2 = 1$ and pick $\bar{\theta}_1$ to match log wage ratio of 90 wage percentile to 10 wage percentile of 4 from Autor et al. (2008). We set the discount factor $\bar{\beta}$ to match an (annual) interest rate of 2%. We set the parameter ψ to match Frisch elasticity of labor supply equal to 0.5. In our model, $\bar{g}$ corresponds to non-transfer government expenditures, which in the U.S. varied from 7% and 11% in the post WWII period and were above 20% during the war. We set $\bar{g}$ to 12% of GDP. Finally, we set Pareto weights α to match the average marginal tax rate in the US of about 20% as in Chari et al. (1994).[21]

Next we turn to the business cycle targets. We calibrate $\{e_{i,\theta}, e_\beta, \Pr(s|s_-)\}$ to match the following four facts about booms and recessions (using NBER dates, for the last 3 recessions i.e. 1991-92, 2001-02 and 2008-10): the log of the incomes individuals at both the 10th and the 90th percentile falls the recessions; 10th percentile income falls by more than 90th percentile; an inflation-adjusted interest rate on government debt is generally lower in recessions; and booms last longer than recessions. We calibrate the average spread in labor productivity to match the average 3% loss in output seen in the last three recessions.The inequality shock is designed to match the facts documented in Guvenen et al. (2012) that the fall in earnings of the 10-percentile is about 2.5 times of 90-percentile. The discount factor shocks match the average boom-recession difference of about 1.96% in the real risk-free interest rate (3 month T bill rate - inflation rate) seen in the last three recessions.[22] We calibrate the transition matrix to get match the average duration of booms and recessions. For comparison, we also report the optimal responses to a drop in government expenditure that leads to an output drop similar magnitude.

Note that because each of them is an exact function of s_t, government expenditures, the discount factor, and productivities are perfectly correlated: a recession is an episode in which TFP falls, inequality

[20] Formally, an equilibrium in an economy where all shocks are forever equal to their mean value

[21] We use federal government expenditures (excluding current transfers) since the labor tax rate of 20% in Chari et al. (1994) is calibrated to federal marginal taxes.

[22] It has long been noticed that the standard RBC model predicts counter-factual negative correlation between real interest rates and output (e.g. Boldrin et al. (2001)). In the data HP filtered output is roughly uncorrelated with real interest rates, but this relationship turn positive if we look at peak vs troughs. We report the optimal responses for both economies with positive and zero correlation of interest rates and output and contrast with a response to a pure TFP shock.

rises, and the discount factor is high. We set the initial level of government debt to be 60%, roughly to match the ratio of federal debt held by public at the beginning of 2010.

Table 1 summarizes some details about our calibration.

Parameter	Value	Description	Target
ψ	0.6994	Frisch elasticity of labor supply	0.5
$\bar{\theta}_1$	4	Log 90-10 wage ratio (Autor et all)	4
$\bar{\theta}_2$	1	Normalize to 1	1
β	0.98	Average (annual) risk free interest rate	2%
$\frac{\hat{\theta}_2}{\hat{\theta}_1}$	2.5	Relative drop in wage income of 10th percentile as compared to 90th percentile	2.5
$\hat{\theta}_1$	1.2%	Average output loss	3%
$\hat{\beta}(s)$	1.96%	Difference in real interest rates between booms and recession	1.96%
α_1	0.69	Marginal tax rate in the economy with no shocks	20%
g	12%	Average pre-transfer expenditure- output ratio	12 %
$P(r\|r)$	0.63	Duration of recessions	2.33 years
$P(b\|b)$	0.84	Duration of booms	7 years

Table 1: Benchmark calibration

7.2 Outcomes

We discuss separately long run and short run implications for optimal policy. In particular, we study the economy (**"Benchmark"**) with the calibration discussed above and a few variants that successively turn off particular sources of variation.

1. **Acyclical Interest Rates**: In the first variant, we recalibrate the discount factor shocks to make the risk-free rate be uncorrelated with output.

2. **Countercyclical Interest Rates**: Here we shut off discount factor shocks by setting $\hat{\beta}(s) = 0$ in (31). Note that under this assumption, interest rates are countercyclical.

3. **No Inequality**: This variant modifies the "Benchmark" by setting $\hat{\beta}(s) = 0$ and $\hat{\theta}_1(s) = \hat{\theta}_2(s) =$ 3% in (31). This corresponds to a case when the only source of business cycle fluctuations is a TFP shock that affects all agents equally. This case more closely matches the experiments in the RBC literature such as Chari et al. (1994).

4. **Government expenditure Shocks**: The last variant compares optimal responses to shocks to government expenditures. In this experiment, we set $\hat{\theta}(s) = \hat{\beta}(s) = 0$ and choose $\hat{g}(s)$ to produce a drop in output of a similar magnitude to that in the first three experiments. This compares to the studies of responses to government shocks by AMSS and Faraglia et al. (2012).

Long run

Figure 6 plots government debt. All experiments start with initial government debt to GDP ratio of 60%. Several features emerge from this figure.

Figure 6: Debt benchmark (o), acyclical interest rates (+), countercyclical interest rates ($\diamond$) and no inequality shocks ($\square$)

In line with Section 6, all four economies the state (x, ρ) converge to some long run ergodic set, so that government debt and the tax rate converge to associated sets. When there are no discount factor shocks (See lines with $\diamond$, $\square$ in figure 6) or small discount factor shocks that produce acyclical interest rates (line with + in figure 6) the government has accumulated assets in this ergodic set. Consistent with 5 the optimal policy adjusts net asset positions to ameliorate the two key constraints impinging on the government policy, namely, the inability to award agent-specific transfers (the restriction to *affine taxes*) and the absence of state-contingent assets (the restriction to *risk-free debt*). Starting from a point when the relative assets of the low skilled agent (or the government if we use the normalization that sets $b_{2,t} = 0$) are low, extracting resources through lower transfers exacerbates inequality. This is costly since the government has to use higher taxes in future to redistribute. On the margin, the optimal policy requires the government (or low skill agent) to accumulate assets. But interest rate fluctuations interact with net asset positions to generate state-contingent earnings from assets. If interest rate are high when the government needs additional revenues, accumulating assets relaxes the restriction imposed by absence of state contingent assets. Thus, with countercyclical interest rates, these forces reinforce each other, making the government's long run asset position be positive.

In data, however, interest rates generally decline in recessions. Procyclical interest rates mean that the two forces outlined in the previous paragraph now oppose each other. For large enough interest rate fluctuations, this means that the government may want to accumulate debt. In Figure 6, the line with (o) represents the benchmark with discount factor shocks rigged to replicate procylical fluctuations in interest rates. For a particular initial condition for government debt, the planner can refrain from varying debt for a very long time.[23]

Convergence to the ergodic region is very slow. With persistent shocks and an initial 60% debt-GDP ratio, it takes about 3,000 years for the government to want to pay off all that debt and then start accumulating assets. With discount factor shocks, it takes even longer to repay the debt. It is still indebted after 5000 years. This confirms the comparative statics of the eigenvalues of linearized system in proposition 6.

Thus, the covariance of interest rates with fundamentals as emphasized in proposition 5 substantially influences the ergodic distribution of government assets.

[23]Like the finding in Proposition 5 for large discount factor shocks (in a way that interest rates are procylical) there exist regions where x_t, ρ_t have low volatility and the government is *not* accumulating assets. But these regions are typically unstable. The two forces highlighted before that guide accumulation of assets now work in opposite direction and the net effect depends on the relative strengths. In particular the sample paths from different initial conditions $\tilde{b}_{2,-1}$ (which would imply different choices for the initial x_0, ρ_0) may display larger fluctuations in assets. However, at the calibrated initial conditions (60% debt-gdp ratio), the uncertainty associated with the mean path is very low for the first 5000 periods.

	Δg	ΔB	ΔT	$\Delta[\tau\theta_1 l_1]$	$\Delta[\tau\theta_2 l_2]$	ΔY	$\Delta\tau$
Benchmark	0.0000	-1.1561	0.6871	-0.1593	-0.3096	-2.8536	0.3732
Acyclical Interest Rates	0.0000	-1.1126	0.6591	-0.1497	-0.3038	-2.8613	0.3879
Countercyclical Interest Rates	0.0000	-1.0794	0.6387	-0.1415	-0.2992	-2.8677	0.3997
No Inequality	0.0000	-0.1380	-0.5459	-0.5635	-0.1204	-2.6294	0.0622
Expenditure Shocks	-7.5037	2.9137	2.8612	-1.3759	-0.3530	-2.3443	-1.1598

Table 2: The tables summarizes the changes in the different components of the government budget as we transit from "boom" to a "recession". All numbers are normalized by un-distorted GDP except τ.

Short run

The analysis of the previous subsection studied aspects of very low frequency components of the optimal policy. Here we focus on business cycle frequencies. In our setting, these higher frequency responses can conveniently be divided into the magnitudes of changes as we switch from "booms" to "recession," and the dynamics during periods when recession or boom state persist.

We set the exogenous state s_0 so that we are in an outset of a recession. Then we solve the time 0 problem with identical initial conditions across different settings. This pins down the initial state vector x_0, ρ_0 that appears in our time 0 Bellman equation. We then use the policy rules to compute fluctuations of different components in the government budget constraint across states. These responses are summarized in Table 2. For each variable z in the table we report in the form $\Delta z \equiv \left(z\left(s_l|x_0,\rho_0,s_0\right) - z\left(s_h|x_0,\rho_0,s_0\right)\right)/\bar{Y}$ where $\bar{Y}$ is average undistorted GDP. [24]

The source of shocks is very important. Three different types of shocks that produce similar drops in GDP have very different consequences for optimal policies, both qualitatively and quantitatively. In the benchmark, the government responds to a shock by a making big increases in transfers, the tax rate, and government debt. However, without inequality shocks (row 4), the government responds by decreasing transfers and increasing both debt and the tax rate, but by an amount an order of magnitude smaller than the benchmark. This indicates that ignoring distributional goals can produce a misleading view about optimal government policy in recessions.

Discount factor shocks have minor effects on impact and matter more for transient dynamics that ultimately have big long run effects. Figures 7 and 8 show how the transient dynamics for prolonged booms (or recessions) differ with and without discount factor shocks. The four panels have taxes, transfers, debt and interest rate movements for a path of 25 years. The bold lines in figures 7 and 8 refer to the benchmark (with procylical interest rates) and the version with acyclical interest rates, respectively. The dotted line in both the figures is the version with countercyclical interest rates. The shaded regions are periods with low output. We see that in a prolonged booms, the government accumulates assets and that it lowers the tax rate when there are no discount factor shocks.

[24] Note that predetermined variables like repayment on existing debt drop out of the accounting and we have

$$\Delta[g] + \Delta[T] + \Delta[B] = \Delta[\tau\theta_1 l_1] + \Delta[\tau\theta_2 l_2]$$

.

Draft25Graphs/DiscoutFactorComparison.png

Figure 7: This plots a typical sample path taxes, transfers, debt and interest rates. The bold lines are with benchmark calibration and the dotted lines refer to the variant with countercyclical interest rates. The shaded regions are recessions.

Draft25Graphs/DiscoutFactorComparison2.png

Figure 8: This plots a typical sample path taxes, transfers, debt and interest rates. The bold lines are with acyclical interest rates calibration and the dotted lines correspond to the case with countercyclical interest rates. The shaded regions are recessions.

8 Concluding remarks

The spring of 2013 witnessed a heated debate in newspapers and economic magazines about the accuracy and meaning of empirical correlations between output growth rates and ratios of government debt to GDP and in data sets assembled by Reinhart and Rogoff (2010). From the perspectives of our paper and of Werning (2007), those correlations and those debates are especially difficult to interpret because in our settings, total government debt is not a relevant state variable that affects allocations, government transfers, or tax rates. The principal message of our paper is that without exogenous restrictions on transfers, the level of government debt is not what matters. What does matter is how government debt is distributed among people relative to society's attitudes toward unequal allocations of consumption and labor. Using a recursive representation that works with a correct state variables — a vector of marginal utility adjusted net asset positions and a vector of pairwise ratios of marginal utilities – we have presented a sequence of examples designed to show how agents net positions affect optimal government policies for choosing distorting tax rates, transfers, and government issues or holdings of risk-free bonds. We find that a significant determinant of an optimal asymptotic government debt or government debt-GDP ratio is how interest rate risk is correlated with risks to fundamentals that threaten to widen or narrow inequality in after-tax and after-transfer incomes. To interpret those Reinhart-Rogoff facts country-by-country, we would want to know much more about the distribution of net assets across people within each country and how they interact with interest rate risks and other risks.

9 Appendix

9.1 Additional details for Section 3.1

In this section we construct an example in which the government can achieve higher welfare in the economy with ad-hoc borrowing limits. We restrict ourselves to $I = 2$, $\theta_1 > \theta_2 = 0$, quasilinear preferences for Agent 1 : $U^1(c,l) = c - h(l)$ and a concave utility function (that satisfies Inada conditions) for Agent 2

Suppose that $g_t = 0$ and $\beta_t = \beta$ for all t, so that the economy is deterministic. In addition, assume that $\underline{b}_1 = 0$ and $\underline{b}_2 = -\infty$. Given $\left(\beta^{-1}b_{1,-1}, \beta^{-1}b_{2,-1}, \beta^{-1}B_{-1}\right)$, the optimal policy solves :

$$\max_{\{c_{1,t},c_{2,t},l_{1,t},b_{1,t},b_{2,t},R_t\}_t} \sum_{t=0}^{\infty} \beta^t \left[\alpha_1 \left(c_{1,t} - h(l_{1,t})\right) + \alpha_2 U^2(c_{2,t})\right] \tag{32}$$

subject to

$$c_{2,t} - c_{1,t} + b_{2,t} - b_{1,t} + h'(l_{1,t})l_{1,t} = R_{t-1}\left(b_{2,t-1} - b_{1,t-1}\right) \tag{33a}$$

$$c_{1,t} + c_{2,t} \le \theta l_{1,t} \tag{33b}$$

$$1 \ge \beta R_t \tag{33c}$$

$$(1 - \beta R_t)b_{1,t} = 0 \tag{33d}$$

$$U^2_{c,t} = \beta R_t U^2_{c,t+1} \tag{33e}$$

$$b_{1,t} \ge 0. \tag{33f}$$

We solve this maximization problem in two stages. First, we solve the problem (32) for a fixed sequence of $\{R_t\}_t$. Denote the value of the objective function for the reduced problem by $W\left(\{R_t\}_t\right)$. Second choose $\{R_t\}_t$ to maximize $W\left(\{R_t\}_t\right)$.

Let $\mu_t\beta^t \Pr(s^t)$ be the Lagrange multiplier associated with the constraint (33a).

Lemma 2 *For $R_t = \beta^{-1}$, we can choose a time invariant solution $c_{1,t} = \bar{c}_1, c_{2,t} = \bar{c}_2, l_{1,t} = \bar{l}_1, b_{1,t} = \bar{b}_1, b_{2,t} = \bar{b}_2$ to the relaxed problem that satisfies*

$$\bar{b}_2 - \bar{b}_1 = \bar{b}_2 = \bar{b}_{2,-1} - \bar{b}_{1,-1}.$$

Proof. By Theorem 1, we can set $b_{1,t} = 0$ and ignore constraint (33f). Further ignoring constraint (33e), a stationary interior solution is given by

$$U^2_c(\bar{c}_2) = \frac{2\bar{\mu} + \alpha_1}{\alpha_2}, \tag{34a}$$

$$\alpha_1 h'(\bar{l}_1) = \theta_1\alpha_1 + \bar{\mu}[\theta_1 - h''(\bar{l}_1)\bar{l}_1 - h'(\bar{l}_1)], \tag{34b}$$

$$b_{2,-1} = \frac{2\bar{c}_2 + \bar{l}_1\left[h'(\bar{l}_1) - \theta_1\right]}{\beta^{-1} - 1}. \tag{34c}$$

Note that if a solution to the above set of equations exists, (33e) is naturally satisfied.

We first establish some comparative statics with respect to $\bar{\mu}$. It is easy to see that concavity of U^2 implies $\frac{\partial \bar{c}_2}{\partial \mu} < 0$. Further equation (34b) can be rearranged to get

$$\frac{h'(\bar{l}_1)}{\theta_1} = \frac{\alpha_1 + \bar{\mu}}{\left[\alpha_1 + \bar{\mu} + \bar{\mu}\left(\frac{h''(\bar{l}_1)\bar{l}_1}{h'(\bar{l}_1)}\right)\right]}$$

.

Using convexity of h we have $\frac{\partial \bar{l}_1}{\partial \mu} < 0$.

As $\bar{\mu} \to -\frac{\alpha_1}{2}$, the RHS of (34c) approaches $+\infty$ and $\bar{\mu} \to +\infty$ it approaches some $\underline{b}_2 < 0$. Thus we have a stationary solution for a range of $b_{2,-1}$.

■

Lemma 3 *For $R_t = \beta^{-1}$, if $b_{2,-1} < b_{1,-1}$ then $\bar{\mu} > 0$.*

Proof. Suppose $\bar{\mu} \leq 0$ when $b_{2,-1} < b_{1,-1}$. At $\mu = 0$, $h'(\bar{l}_1) = \theta_1$ and the RHS of (34c) is positive. At $\mu < 0$ we have $h'(\bar{l}_1) > \theta_1$. The observations above imply that the RHS of (34c) is increasing in μ and this clearly violates equation (34c). Thus we have a contradiction. ■

When $R_t = 1/\beta$ for all t, the solution of the reduced problem is an optimal allocation for an economy in which agents face no borrowing constraints

Let $\frac{\partial}{\partial R_1} W(\{R_t\}_t)\Big|_{\{R_t\}=\boldsymbol{\beta}^{-1}}$ be the derivative of $W(\{R_t\}_t)$ with respect to R_1 evaluated at $R_t = \beta^{-1}$ for all t. The multiplier on constraints (33d) and (33e) are zero and let $\xi_t \geq 0$ be the multiplier on constraint (33c). Our observations above imply that

$$\frac{\partial}{\partial R_1} W(\{R_t\}_t)\Big|_{\{R_t\}=\boldsymbol{\beta}^{-1}} = \bar{\mu}\bar{b}_{2,t} - \xi_1 \leq \bar{\mu}\left(\bar{b}_{2,-1} - \bar{b}_{1,-1}\right) < 0,$$

and therefore $R_t = \beta^{-1}$ for all t is not the optimal equilibrium sequence. Therefore, welfare in the economy with exogenous borrowing constraints is strictly higher than in the economy without exogenous borrowing constraints.[25]

The outcome that welfare can be strictly higher with exogenous borrowing constraints depends on our assumption that agents do not face idiosyncratic risk. If agents were also subject to idiosyncratic shocks, exogenous borrowing constraints would have the additional effect of limiting agents' ability to self-insure against those shocks.[26] Nevertheless, the insight from the example carries through that even though exogenous borrowing constraints can hurt agents' to insure against idiosyncratic shocks, they can help a government smooth distortions with respect to aggregate shocks like government expenditure shocks.

[25] The mechanism in this example is similar to a finding of **?**, who showed that relaxing agents' borrowing constraints can be suboptimal in an economy with idiosyncratic shocks. Our analysis shows that this insight is more general and holds even in economies with no shocks.

[26] See Aiyagari and McGrattan (1998) and Heathcote (2005) for details.

9.2 Proof of Proposition 4

We prove a slight more general version of our result. Consider an infinite horizon, incomplete markets economy in which an agent maximizes utility function $U : \mathbb{R}^n_+ \to \mathbb{R}$ subject to an infinite sequence of budget constraints. We assume that U is concave and differentiable. Let $\boldsymbol{a}(s^t)$ be a vector of n goods and let $\boldsymbol{p}(s^t)$ be a price vector in state s^t with $p_i(s^t)$ denoting the price of good i. We use a normalization $p_1\left(s^t\right) = 1$ for all s^t. There is a risk-free bond.

Let $b(s^t)$ be the agent's bond holdings, and let $\boldsymbol{e}\left(s^t\right)$ be a stochastic vector of endowments.

Consumer maximization problem

$$\max_{\boldsymbol{a}_t, b_t} \sum_{t=0}^{\infty} \left[\Pi_{j=0}^{t} \beta(s_j)\right] \Pr\left(s^t\right) U(\boldsymbol{a}\left(s^t\right)) \tag{35}$$

subject to

$$\boldsymbol{p}\left(s^t\right) \boldsymbol{a}\left(s^t\right) + q(s^t) b\left(s^t\right) = \boldsymbol{p}\left(s^t\right) \boldsymbol{e}\left(s^t\right) + b\left(s^{t-1}\right) \tag{36}$$

and $\left\{b\left(s^t\right)\right\}$ is bounded and $\left\{q(s^t)\right\}$ is the price of the risk-free bond.

The Euler conditions are

$$\begin{aligned} \boldsymbol{U}_a(s^t) &= U_1(s^t)\boldsymbol{p}(s^t) \\ \Pr\left(s^t\right) U_1\left(s^t\right) q(s^t) &= \beta(s_t) \sum_{s^{t+1} > s^t} \Pr\left(s^{t+1}\right) U_1\left(s^{t+1}\right). \end{aligned} \tag{37}$$

Lemma 4 *Consider an allocation $\{\boldsymbol{a}_t, b_t\}$ that satisfies (36), (37) and $\{b_t\}_t$ is bounded. Then $\{\boldsymbol{a}_t, b_t\}$ is a solution to (35).*

Proof. The proof follows closely Constantinides and Duffie (1996). Suppose there is another budget feasible allocation $\boldsymbol{a} + \boldsymbol{h}$ that maximizes (35). Since U is strictly concave,

$$\begin{aligned} & \mathbb{E}_0 \sum_{t=0}^{\infty} \left[\Pi_{j=0}^{t} \beta(s_j)\right] U(\boldsymbol{a}_t + \boldsymbol{h}_t) - \mathbb{E}_0 \sum_{t=0}^{\infty} \left[\Pi_{j=0}^{t} \beta(s_j)\right] U(\boldsymbol{a}_t) \\ \leq\ & \mathbb{E}_0 \sum_{t=0}^{\infty} \left[\Pi_{j=0}^{t} \beta(s_j)\right] \boldsymbol{U}_a(\boldsymbol{a}_t) \boldsymbol{h}_t \end{aligned} \tag{38}$$

To attain $\boldsymbol{a}+\boldsymbol{h}$, the agent must deviate by φ_t from his original portfolio b_t such that $\{\varphi_t\}_t$ is bounded, $\varphi_{-1} = 0$ and

$$\boldsymbol{p}(s^t)\boldsymbol{h}\left(s^t\right) = \varphi(s^{t-1}) - q(s^t)\varphi(s^t)$$

Multiply by $\left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \Pr\left(s^t\right) U_1(s^t)$ to get:

$$\begin{aligned} \left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \Pr\left(s^t\right) U_1(s^t) \boldsymbol{p}(s^t) \boldsymbol{h}\left(s^t\right) &= \left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \Pr\left(s^t\right) U_1(s^t) \varphi(s^{t-1}) - q(s^t) \left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \Pr\left(s^t\right) U_1(s^t) \varphi(s^t) \\ &= \left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \Pr\left(s^t\right) U_1(s^t) \varphi(s^{t-1}) - \left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \beta(s_t) \sum_{s^{t+1} > s^t} \Pr\left(s^{t+1}\right) U_1\left(s^{t+1}\right) \varphi(s^t) \end{aligned}$$

where we used the second part of (37) in the second equality. Sum over the first T periods (pathwise) and use the first part of (37) to eliminate $\boldsymbol{U}_a(\boldsymbol{a}_t) = U_1(s^t)\boldsymbol{p}(s^t)$

$$\sum_{t=0}^{T} \left[\Pi_{j=0}^{t-1} \beta(s_j)\right] \Pr\left(s^t\right) \boldsymbol{U}_a(\boldsymbol{a}_t) \boldsymbol{h}\left(s^t\right) = -\left[\Pi_{j=0}^{T} \beta(s_j)\right] \sum_{s^{T+1} > s^T} \Pr\left(s^{T+1}\right) U_1\left(s^{T+1}\right) \varphi(s^T).$$

Since $\{\varphi_t\}_t$ is bounded there must exist $\bar{\varphi}$ s.t. $|\varphi_t| \leq \bar{\varphi}$ for all t. By Theorem 5.2 of Magill and Quinzii (1994), this equilibrium with debt constraints implies a transversality condition on the right hand side of the last equation, so by transitivity we have

$$\lim_{T\to\infty} \sum_{t=0}^{T} \left[\Pi_{j=0}^{t-1}\beta(s_j)\right] \Pr\left(s^t\right) \boldsymbol{U}_a(\boldsymbol{a}_t)\boldsymbol{h}\left(s^t\right) = 0.$$

Substitute this into (38) to show that $\boldsymbol{h}$ does not improve utility of consumer. ■

9.3 Additional details for Lemma 1

Given $P(s)\mu_i(s)$ and λ_i be the multipliers on constraints (24) and (25). The first order conditions are then as follows

$x_i'(s)$:

$$\beta(s)V_{x_i}(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s)) - \beta(s)\mu_i(s) = 0 \tag{39}$$

$\tau(s)$:

$$\begin{aligned} &U\tau(\tau(s), \boldsymbol{\rho}'(s), s) + \sum_i \left(\frac{x_i u_{c,N\tau}(\tau(s), \boldsymbol{\rho}'(s), s)}{\mathbb{E}u_{c,1}(\tau, \boldsymbol{\rho}', s)} \left[\mu_i(s) - \frac{\mathbb{E}\mu_i(s)u_{c,1}(\tau, \boldsymbol{\rho}', s)}{\mathbb{E}u_{c,1}(\tau, \boldsymbol{\rho}', s)}\right] - \mu_i(s) Z_{i,\tau}(\tau(s), \boldsymbol{\rho}'(s), s) \right) \\ &\quad + \sum_i \left[\lambda_i u_{c,i\tau}(\tau(s), \boldsymbol{\rho}'(s), s)(\rho_i'(s) - \rho_i)\right] = 0 \end{aligned} \tag{40}$$

$\rho_i(s)$:

$$\begin{aligned} &U_{\rho_i}(\tau(s), \boldsymbol{\rho}'(s), s) + \sum_j \left(\frac{x_j u_{c,1\rho_i}(\tau(s), \boldsymbol{\rho}'(s), s)}{\mathbb{E}u_{c,1}(\tau, \boldsymbol{\rho}', s)} \left[\mu_j(s) - \frac{\mathbb{E}\mu_j(s)u_{c,1}(\tau, \boldsymbol{\rho}', s)}{\mathbb{E}u_{c,1}(\tau, \boldsymbol{\rho}', s)}\right] - \mu_j(s) Z_{j,\rho_i}(\tau(s), \boldsymbol{\rho}'(s), s) \right) \\ &\quad + \sum_j \left[\lambda_j u_{c,j\rho_i}(\tau(s), \boldsymbol{\rho}'(s), s)(\rho_j'(s) - \rho_j)\right] + \lambda_i u_{c,i}(\tau(s), \boldsymbol{\rho}'(s), s) + \beta(s)V_{\rho_i}(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s)) = 0. \end{aligned} \tag{41}$$

Finally the envelope conditions are

$$V_{x_i}(\boldsymbol{x}, \boldsymbol{\rho}) = \frac{\mathbb{E}\mu_i(s)u_{c,1}(\tau, \boldsymbol{\rho}', s)}{\mathbb{E}u_{c,1}(\tau, \boldsymbol{\rho}', s)}, \tag{42}$$

and

$$V_{\rho_i}(\boldsymbol{x}, \boldsymbol{\rho}) = -\lambda_i \mathbb{E}u_{c,i}(\tau, \boldsymbol{\rho}', s). \tag{43}$$

The equations for the steady state can then be obtained by imposing $x_i'(s) = x_i$ and $\rho_i'(s) = \rho_i$. It is then readily noted that equations (39) and (42) are satisfied when $\mu_i(s) = \mu_i = \beta V_{x_i}(\boldsymbol{x}, \boldsymbol{\rho})$. Further equation (25) drops out and the equation (24) simplifies to

$$Z_i(\tau(s), \boldsymbol{\rho}, s) + \beta(s)x_i' = \frac{x_i U_c^1(\tau(s), \boldsymbol{\rho}, s)}{\mathbb{E}U_c^1(\tau, \rho)}.$$

9.4 Proof of Proposition 5

The Bellman equation for the optimal planners problem with log quadratic preferences and IID shocks can be written as

$$V(x,\rho) = \max_{c_1,c_2,l_1,x',\rho'} \sum_s \Pr(s)\left[\alpha_1\left(\log c_1(s) - \frac{l_1(s)^2}{2}\right) + \alpha_2 \log c_2(s) + \beta(s)V(x'(s),\rho'(s))\right]$$

subject to the constraints

$$1 + \rho'(s)[l_1(s)^2 - 1] + \beta(s)x'(s) - \frac{x\frac{1}{c_2(s)}}{\mathbb{E}[\frac{1}{c_2}]} = 0 \tag{44}$$

$$\sum_s \frac{\Pr(s)}{c_1(s)}(\rho'(s) - \rho) = 0 \tag{45}$$

$$\theta_1(s)l_1(s) - c_1(s) - c_2(s) - g = 0 \tag{46}$$

$$\rho'(s)c_2(s) - c_1(s) = 0 \tag{47}$$

where the $\Pr(s)$ is the probability distribution of the aggregate state s. If we let $\Pr(s)\mu(s)$, λ, $\Pr(s)\xi(s)$ and $\Pr(s)\phi(s)$ be the Lagrange multipliers for the constraints (44)-(47) respectively then we obtain the following FONC for the planners problem

$c_1(s)$:

$$\frac{\alpha_1 \Pr(s)}{c_1(s)} - \frac{\lambda \Pr(s)}{c_1(s)^2}(\rho'(s) - \rho) - \Pr(s)\xi(s) - \Pr(s)\phi(s) = 0 \tag{48}$$

$c_2(s)$:

$$\frac{\alpha_2 \Pr(s)}{c_2(s)} + \frac{x\Pr(s)}{c_2(s)^2\mathbb{E}[\frac{1}{c_2}]}\left[\mu(s) - \frac{\mathbb{E}[\mu\frac{1}{c_2}]}{\mathbb{E}[\frac{1}{c_2}]}\right] - \Pr(s)\xi(s) + \Pr(s)\rho'(s)\phi(s) = 0 \tag{49}$$

$l_1(s)$:

$$-\alpha_1 \Pr(s)l_1(s) + 2\mu(s)\Pr(s)\rho'(s)l_1(s) + \theta_1(s)\Pr(s)\xi(s) = 0 \tag{50}$$

$x'(s)$:

$$\beta(s)\Pr(s)V_x(x'(s),\rho'(s)) + \beta(s)\Pr(s)\mu(s) = 0 \tag{51}$$

$\rho'(s)$:

$$\beta(s)\Pr(s)V_\rho(x'(s),\rho'(s)) + \frac{\lambda\Pr(s)}{c_1(s)} + \mu(s)\Pr(s)[l_1(s)^2 - 1] + \Pr(s)\phi(s)c_2(s) = 0 \tag{52}$$

In addition there are two envelope conditions given by

$$V_x(x,\rho) = -\sum_{s'} \frac{\mu(s')\Pr(s')\frac{1}{c_2(s')}}{\mathbb{E}[\frac{1}{c_2}]} = -\frac{\mathbb{E}[\mu\frac{1}{c_2}]}{\mathbb{E}[\frac{1}{c_2}]} \tag{53}$$

$$V_\rho(x,\rho) = -\lambda\mathbb{E}[\frac{1}{c_1}] \tag{54}$$

A steady state is then a collection of allocations, initial conditions and Lagrange multipliers $\{c_1(s), c_2(s), l_1(s), x, \rho, \mu(s), \lambda, \xi(s), \phi(s)\}$ such that equations (44)-(54) are satisfied when $\rho'(s) = \rho$ and

$x'(s) = x$. It should be clear that is that if we replace $\mu(s) = \mu$ then, equation (51) is always satisfied. Additionally under this assumption equation (49) simplifies significantly.

$$\frac{x \Pr(s)}{c_2(s)^2 \mathbb{E}[\frac{1}{c_2}]} \left[\mu(s) - \frac{\mathbb{E}[\mu \frac{1}{c_2}]}{\mathbb{E}[\frac{1}{c_2}]} \right] = 0$$

The first order conditions for a steady can then be written simply as

$$1 + \rho[l_1(s)^2 - 1] + \beta(s)x - \frac{x}{c_2(s)\mathbb{E}[\frac{1}{c_2}]} = 0 \tag{55}$$

$$\theta_1(s)l_1(s) - c_1(s) - c_2(s) - g = 0 \tag{56}$$

$$\rho c_2(s) - c_1(s) = 0 \tag{57}$$

$$\frac{\alpha_1}{c_1(s)} - \xi(s) - \phi(s) = 0 \tag{58}$$

$$\frac{\alpha_2}{c_2(s)} - \xi(s) + \rho\phi(s) = 0 \tag{59}$$

$$[2\mu\rho - \alpha_1]l_1(s) + \theta_1(s)\xi(s) = 0 \tag{60}$$

$$\lambda \left[\frac{1}{c_1(s)} - \beta(s)\mathbb{E}[\frac{1}{c_1}] \right] + \mu[l_1(s)^2 - 1] + \phi(s)c_2(s) = 0 \tag{61}$$

We can rewrite equation (58) as

$$\frac{\alpha_1}{c_2(s)} - \rho\xi(s) - \rho\phi(s) = 0$$

by substituting $c_1(s) = \rho c_2(s)$. Adding this to equation (59) and normalizing $\alpha_1 + \alpha_2 = 1$ we obtain

$$\xi(s) = \frac{1}{(1+\rho)\, c_2(s)} \tag{62}$$

which we can use to solve for $\phi(s)$ as

$$\phi(s) = \frac{\alpha_1 - \rho\alpha_2}{(\rho(1+\rho))\, c_2(s)} \tag{63}$$

From equation (55) we can solve for $l_1(s)^2 - 1$ as

$$l_1(s)^2 - 1 = \frac{x}{\rho \mathbb{E}[\frac{1}{c_2}]} \left(\frac{1}{c_2(s}) - \beta(s)\mathbb{E}[\frac{1}{c_2}] \right) - \frac{1}{\rho}$$

This can be used along with equations (61) and (63) to obtain

$$\left(\frac{\lambda}{\rho} + \frac{\mu x}{\rho \mathbb{E}[\frac{1}{c_2}]} \right) \left(\frac{1}{c_2(s)} - \beta(s)\mathbb{E}\left[\frac{1}{c_2}\right] \right) = \frac{\mu}{\rho} + \frac{\rho\alpha_2 - \alpha_1}{\rho(1+\rho)}$$

The LHS depends on s while the RHS does not, Hence the solution to this equation is

$$\lambda = -\frac{\mu x}{\mathbb{E}[\frac{1}{c_2}]} \tag{64}$$

and

$$\mu = \frac{\alpha_1 - \rho\alpha_2}{1+\rho} \tag{65}$$

Combining these with equation (60) we quickly obtain that

$$\left[2\rho\frac{\alpha_1-\rho\alpha_2}{1+\rho}-\alpha_1\right]l_1(s)+\frac{\theta_1(s)}{(1+\rho)\,c_2(s)}=0$$

Then solving for $l_1(s)$ gives

$$l_1(s)=\frac{\theta_1(s)}{(\alpha_1(1-\rho)+2\rho^2\alpha_2)\,c_2(s)}$$

Remark 1 *Note that the labor tax rate is given by $1-\frac{c_1(s)l_1(s)}{\theta(s)}$. The previous expression shows that labor taxes are constant at the steady state. This property holds generally for CES preferences separable in consumption and leisure*

This we can plug into the aggregate resource constraint (56) to obtain

$$l_1(s)=\left(\frac{1+\rho}{\alpha_1(1-\rho)+2\rho^2\alpha_2}\right)\frac{1}{l_1(s)}+\frac{g}{\theta_1(s)}$$

letting $C(\rho)=\frac{1+\rho}{\alpha_1(1-\rho)+2\rho^2\alpha_2}$ we can then solve for $l_1(s)$ as

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

The marginal utility of agent 2 is then

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

Note that in order for either of these terms to be positive we need $C(\rho)\geq 0$ implying that there is only one economically meaningful root. Thus

$$l_1(s)=\frac{g+\sqrt{g^2+4C(\rho)\theta_1(s)^2}}{2\theta_1(s)} \tag{66}$$

and

$$\frac{1}{c_2(s)}=\left(\frac{1+\rho}{C(\rho)}\right)\left(\frac{g+\sqrt{g^2+4C(\rho)\theta_1(s)^2}}{2\theta_1(s)^2}\right) \tag{67}$$

A steady state is then a value of ρ such that

$$x(s)=\frac{1+\rho[l_1(\rho,s)^2-1]}{\frac{1/c_2(\rho,s)}{\mathbb{E}[\frac{1}{c_2}](\rho)}-\beta(s)} \tag{68}$$

s independent of s.

The following lemma, which orders consumption and labor across states, will be useful in proving the parts of proposition 5. As a notational aside we will often use $\theta_{1,l}$ and $\theta_{1,h}$ to refer to $\theta_1(s_l)$ and $\theta_1(s_h)$ respectively. Where s_l refers to the low TFP state and s_h refers to the high TFP state.

Lemma 5 *Suppose that $\theta_1(s_l)<\theta_2(s_h)$ and ρ such that $C(\rho)>0$ then*

$$l_{1,l}=\frac{g+\sqrt{g^2+4C(\rho)\theta_{1,l}^2}}{2\theta_{1,l}}>\frac{g+\sqrt{g^2+4C(\rho)\theta_{1,h}^2}}{2\theta_{1,h}}=l_{1,h}$$

and

$$\frac{1}{c_{2,l}} = \frac{1+\rho}{C(\rho)}\frac{g+\sqrt{g^2+4C(\rho)\theta_{1,l}^2}}{2\theta_{1,l}^2} > \frac{1+\rho}{C(\rho)}\frac{g+\sqrt{g^2+4C(\rho)\theta_{1,h}^2}}{2\theta_{1,h}^2} = \frac{1}{c_{2.h}}$$

Proof. The results should follow directly from showing that the function

$$l_1(\theta) = \frac{g+\sqrt{g^2+4C(\rho)\theta}}{2\theta}$$

is decreasing in θ. Taking the derivative with respect to θ

$$\begin{aligned}
\frac{dl_1}{d\theta}(\theta) &= -\frac{g}{2\theta^2} - \frac{\sqrt{g+4C(\rho)\theta^2}}{2\theta^2} + \frac{4C(\rho)\theta}{2\theta\sqrt{g^2+4C(\rho)\theta^2}} \\
&= -\frac{g}{2\theta^2} - \frac{g+4C(\rho)\theta^2-4C(\rho)\theta^2}{2\theta^2\sqrt{g^2+4C(\rho)\theta^2}} \\
&= -\frac{g}{2\theta^2} - \frac{g}{2\theta^2\sqrt{g^2+4C(\rho)\theta^2}} < 0
\end{aligned}$$

That $\frac{1}{c_{2,l}} > \frac{1}{c_{2,h}}$ follows directly. ■

Proof of Proposition 5.

Part 1. In order for there to exist a ρ such that equation (68) is independent of the state (and hence have a steady state) we need the existence of root for the following function

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

From lemma 5 we can conclude that

$$1+\rho[l_1(\rho,s_l)^2-1] > 1+\rho[l_1(\rho,s_h)^2-1] \tag{69}$$

and

$$\frac{1/c_2(\rho,s_l)}{\mathbb{E}[\frac{1}{c_2}](\rho)}-\beta > \frac{1/c_2(\rho,s_h)}{\mathbb{E}[\frac{1}{c_2}](\rho)}-\beta \tag{70}$$

for all $\rho > 0$ such that $C(\rho) \geq 0$. To begin with we will define $\underline{\rho}$ such that $C(\rho) > 0$ for all $\rho > \underline{\rho}$. Note that we will have to deal with two different cases.

$\alpha_1(1-\rho)+2\rho^2\alpha_2 > 0$ **for all** $\rho \geq 0$**:** In this case we know that $C(\rho) \geq 0$ for all ρ and is bounded above and thus we will let $\underline{\rho} = 0$.

$\alpha_1(1-\rho)+2\rho^2\alpha_2 = 0$ **for some** $\rho > 0$**:** In this case let $\underline{\rho}$ be the largest positive root of $\alpha_1(1-\rho)+2\rho^2\alpha_2$. Note that $\lim_{\rho\to\underline{\rho}^+} C(\rho) = \infty$

With this we note that[27]

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

[27]In the first case $\underline{\rho} = 0$ and in the second case $l_1(\rho,s_l) = l_1(\rho,s_h)$ as $\rho \to \underline{\rho}^+$

We can also show that

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

which implies that $\lim_{\rho\to\underline{\rho}^+} f(\underline{\rho}) > 0$.

Taking the limit as $\rho \to \infty$ we see that $C(\rho) \to 0$, given that $\frac{g}{\theta(s)} < 1$, we can then conclude that

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

Thus, there exists $\overline{\rho}$ such that $1 + \overline{\rho}[l_1(\overline{\rho}, s_l)^2 - 1] = 0$. [28]. From equation (69), we know that

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

which implies in the limit

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

which along with

$$\frac{\frac{1/c_2(\rho,s_h)}{\mathbb{E}[\frac{1}{c_2}]} - \beta}{\frac{1/c_2(\rho,s_l)}{\mathbb{E}[\frac{1}{c_2}]} - \beta} \geq -1$$

allows us to conclude that $\lim_{\rho\to\overline{\rho}^-} f(\rho) = -\infty$. The intermediate value theorem then implies that there exists ρ_{SS} such that $f(\rho_{SS}) = 0$ and hence that ρ_{SS} is a steady state.

Finally, as $\rho_{SS} < \overline{\rho}$ we know that

$$1 + \rho_{SS}[l_1(\rho_{SS}, s_l) - 1] > 0$$

as $\frac{1/c_2(\rho,s_l)}{\mathbb{E}[\frac{1}{c_2}]} > 1$ we can conclude

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

implying that the government will hold assets in the steady state (under the normalization that agent 2 holds no assets).

Part 2. The condition that $R(s_h) = R(s_l)$ implies that

$$\frac{1/c_2(\rho, s_l)}{\beta(s_l)\mathbb{E}[\frac{1}{c_2}]} = \frac{1/c_2(\rho, s_h)}{\beta(s_h)\mathbb{E}[\frac{1}{c_2}]}$$

which simplifies to

$$\frac{\beta(s_h)}{\beta(s_l)} = \frac{1/c_2(\rho, s_h)}{1/c_2(\rho, s_l)} \tag{71}$$

[28]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 $(\underline{\rho}, \infty)$

In order for a steady state to exist with constant interest rates there must be a root of the following function

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

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

$$= \frac{1+\rho[l_1(\rho,s_h)^2-1]}{1+\rho[l_1(\rho,s_l)^2-1]} - \frac{1/c_2(\rho,s_h)}{1/c_2(\rho,s_l)}$$

Taking limits of $f(\rho)$ as ρ approaches $\underline{\rho}$ from the positive side we already demonstrated

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

From equation (67) and Lemma 5 it is straightforward to see that

$$\lim_{\rho\to\underline{\rho}^+} \frac{1/c_2(\rho,s_h)}{1/c_2(\rho,s_l)} < 1$$

which allows us to conclude that

$$\lim_{\rho\to\underline{\rho}^+} f(\rho) > 0$$

Taking limits as ρ approaches $\overline{\rho}$ from the negative direction we know that

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

As $\frac{1/c_2(\rho,s_h)}{1/c_2(\rho,s_h)} > 0$ for all ρ it is straightforward to conclude that

$$\lim_{\rho\to\overline{\rho}^-} f(\rho) = -\infty$$

Continuity then implies the existence of a ρ^{SS} such that $f(\rho^{SS}) = 0$, and thus there exists a $\beta(s_l)$ and $\beta(s_h)$ such that $R(s_l) = R(s_h)$ in steady state. From Lemma 5

$$l(\rho,s_l) > l(\rho,s_h).$$

In order for

$$\frac{1+\rho^{SS}[l_1(\rho^{SS},s_h)^2-1]}{1+\rho^{SS}[l_1(\rho^{SS},s_l)^2-1]} = \frac{1/c_2(\rho^{SS},s_h)}{1/c_2(\rho^{SS},s_l)} < 1$$

it is necessary that

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

implying that the steady state asset level

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

Part 3 As noted before, since $g/\theta(s) < 1$ for all s we have

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

Thus, there exists ρ_{SS} such that

$$0 > 1 + \rho_{SS}[l_1(\rho_{SS}, s_l)^2 - 1] > 1\rho_{SS}[l_1(\rho_{SS}, s_h)^2 - 1]$$

It is then possible to choose $\beta(s) < \frac{1/c_2(\rho_{SS},s)}{\mathbb{E}[\frac{1}{c_2}]}$ such that

$$1 > \frac{1 + \rho_{SS}[l_1(\rho_{SS}, s_l)^2 - 1]}{1 + \rho_{SS}[l_1(\rho_{SS}, s_h)^2 - 1]} = \frac{\frac{1/c_2(\rho_{SS},s_l)}{\mathbb{E}[\frac{1}{c_2}]} - \beta(s_l)}{\frac{1/c_2(\rho_{SS},s_h)}{\mathbb{E}[\frac{1}{c_2}]} - \beta(s_h)} \tag{72}$$

Implying that for discount factor shocks $\beta(s)$, ρ_{SS} is a steady state level for the ratio of marginal utilities, with steady state marginal utility weighted government debt

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

Thus, in the steady state, the government is holding debt, under the normalization that the unproductive worker holds no assets. As $\frac{1/c_2(\rho,s_l)}{\mathbb{E}[\frac{1}{c_2}]} > \frac{1/c_2(\rho,s_h)}{\mathbb{E}[\frac{1}{c_2}]}$, in order for equation (72) to hold we need $\beta_l > \beta_h$. We can then rewrite equation (72) as

$$1 > \frac{\beta(s_h)}{\beta(s_l)} > \frac{\frac{1/c_2(\rho_{SS},s_l)}{\beta_l\mathbb{E}[\frac{1}{c_2}]} - 1}{\frac{1/c_2(\rho_{SS},s_h)}{\beta(s_h)\mathbb{E}[\frac{1}{c_2}]} - 1}$$

Thus

$$R(s_l)\frac{1/c_2(\rho_{SS}, s_l)}{\beta(s_l)\mathbb{E}[\frac{1}{c_2}]} < \frac{1/c_2(\rho_{SS}, s_h)}{\beta(s_h)\mathbb{E}[\frac{1}{c_2}]} = R(s_h) \tag{73}$$

in the steady state interest rates are positively correlated with TFP.

■

9.5 Linearization Algorithm

This section will outline our numerical methods used to solve for and linearize around the steady state in the case of a 2 state iid process for the aggregate state.

$$V(\boldsymbol{x}, \boldsymbol{\rho}) = \max_{c_i(s), l_i(s), \boldsymbol{x}'(s), \boldsymbol{\rho}'(s)} \sum_s P(s) \left(\left[\sum_i \pi_i \alpha_i U(c_i(s), l_i(s)) \right] + \beta(s) V(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s)) \right) \tag{74}$$

$$U_{c,i}(s)c_i(s) + U_{l,i}(s)l_i(s) - \rho_i'(s)\left[U_{c,1}(s)c_1(s) + U_{l,1}(s)l_1(s)\right] + \beta(s)x_i'(s) = \frac{x_i U_{c,i}(s)}{\mathbb{E}U_{c,i}} \tag{75a}$$

$$\sum_s P(s)U_{c,1}(s)(\rho_i(s) - \rho_i) = 0 \tag{75b}$$

$$\frac{\rho'(s)}{\theta_1(s)}U_{l,1}(s) = \frac{1}{\theta_i(s)}U_{l,i}(s) \tag{75c}$$

$$\sum_{j=0}^{N} \pi_j c_j(s) + g(s) = \sum_{j=0}^{N} \pi_j \theta_j(s) l_j(s) \tag{75d}$$

$$U_{c,i}(s) = \rho_i'(s)U_{c,1}(s) \tag{75e}$$

For $i = 2, \ldots, N$. Note that some of the constraints have been modified a little for ease of differentiation. Associated with these constraints we have the Lagrange multipliers $P(s)\mu_i'(s)$, λ_i,$P(s)\phi_i(s)$,$\Pr(s)\xi(s)$, and $P(s)\zeta_i(s)$.

The first order conditions with respect to the choice variables are as follows (note we will be using the notation $\mathbb{E}z$ to represent $\sum_s \Pr(s)z(s)$ for some variable z)

$c_1(s)$**:**

$$\begin{aligned}\pi_1\alpha_1 U_{c,1}(s) &+ \sum_{i=2}^{N}\left(\mu_i'(s)\rho_i'(s)\right)\left[U_{cc,1}(s)c_1(s) + U_{c,1}(s)\right] \\ &+ \lambda U_{cc,1}(s)\sum_{i=2}^{N}(\rho_i'(s) - \rho_i) - \pi_1\xi(s) + \sum_{i=2}^{N}\zeta_i(s)\rho_i'(s)U_{cc,1}(s) = 0\end{aligned} \tag{76a}$$

$c_i(s)$**:** for $i \geq 2$

$$\pi_i\alpha_i U_{c,i}(s) - \mu_i'(s)\left[U_{cc,i}(s)c_i(s) + U_{c,i}(s)\right] + \frac{x_i U_{cc,i}(s)}{\mathbb{E}U_{c,i}}\left(\mu_i'(s) - \frac{\mathbb{E}\mu_i' U_{c,i}}{\mathbb{E}U_{c,i}}\right) - \pi_i\xi(s) - \zeta_i(s)U_{cc,i}(s) = 0 \tag{76b}$$

$l_1(s)$**:**

$$\pi_1\alpha_1 U_{l,1}(s) + \sum_{i=2}^{N}\mu_i'(s)\rho_i(s)\left[U_{ll,1}(s)l_1(s) + U_{l,1}(s)\right] - \sum_{i=2}^{N}\frac{\rho_i'(s)\phi_i(s)}{\theta_1(s)}U_{ll,1}(s) + \pi_1\theta_1(s)\xi(s) = 0 \tag{76c}$$

$l_2(s)$**:**

$$\pi_i\alpha_i U_{l,i}(s) - \mu_i'(s)\left[U_{ll,i}(s)l_i(s) + U_{l,i}(s)\right] + \frac{\phi_i(s)}{\theta_i(s)}U_{ll,i}(s) + \pi_i\theta_i(s)\xi(s) = 0 \tag{76d}$$

$\rho_i'(s)$**:**

$$\beta(s)V_{\rho_i}(\boldsymbol{x}'(s), \boldsymbol{\rho}_i'(s)) + \mu_i'(s)\left[U_{c,1}(s)c_1(s) + U_{l,1}(s)l_1(s)\right] + \lambda_i U_{c,1}(s) - \phi_i(s)\frac{U_{l,1}(s)}{\theta_1(s)} + U_{c,1}(s)\zeta_i(s) = 0 \tag{76e}$$

$x_i'(s)$**:**

$$V_{x_i}(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s)) - \mu_i'(s) = 0. \tag{76f}$$

Equations (75a)-(75e) and (76a)-(76e) then define the necessary conditions for an interior maximization of the planners problem for the state (x, ρ). In addition to these we have the two envelop conditions

$$V_{x_i}(\boldsymbol{x}, \boldsymbol{\rho}) = \frac{\sum_s P(s)\mu_i'(s)U_{c,i}(s)}{\mathbb{E}U_{c,i}(s)} = \frac{\mathbb{E}\mu_i' U_{c,i}}{\mathbb{E}U_{c,i}}, \tag{77a}$$

and

$$V_{\rho_i}(\boldsymbol{x}, \boldsymbol{\rho}) = -\lambda_i \mathbb{E}U_{c,1}. \tag{77b}$$

In order to check local stability we linearize locally around the steady state. Furthermore we find that the policy functions have better numerical properties when the state variables are chosen to be $\boldsymbol{\mu}, \boldsymbol{\rho}$ rather than $\boldsymbol{x}, \boldsymbol{\rho}$, and thus, we will proceed with the linearization procedure using $(\boldsymbol{\mu}, \boldsymbol{\rho})$ as the endogenous state vector. The evolution of the state variable μ must follow the weighted martingale

$$\mu_i - \frac{\sum_s P(s)\mu_i'(s)U_{c,i}(s)}{\sum_s P(s)U_{c,i}(s)} = 0. \tag{78}$$

The optimal policy function, which we will denote as $z(\boldsymbol{\mu}, \boldsymbol{\rho})$, must satisfy $F(z, y, g(z)) = 0$ where F represents the system of equations (75a)-(76e)and (78), y is the state vector $(\boldsymbol{x}, \boldsymbol{\rho})$, and g is the mapping of the policies into functions of future variables, namely $\boldsymbol{x}'(s)$ and $V_{\boldsymbol{\rho}}(\boldsymbol{\mu}'(s), \boldsymbol{\rho}(s))$. In other words

$$g(z) = \begin{pmatrix} \boldsymbol{x}(\boldsymbol{\mu}'(1), \boldsymbol{\rho}'(1)) \\ V_{\boldsymbol{\rho}}(\boldsymbol{\mu}'(1), \boldsymbol{\rho}'(1) \\ \boldsymbol{x}(\boldsymbol{\mu}'(2), \boldsymbol{\rho}'(2)) \\ V_{\boldsymbol{\rho}}(\boldsymbol{\mu}'(2), \boldsymbol{\rho}'(2)) \end{pmatrix}.$$

Finally $z(\boldsymbol{\mu}, \boldsymbol{\rho})$ are the stacked variables $\{c_1(s), c_i(s), l_1(s), l_i(s), \boldsymbol{x}, \boldsymbol{\rho}'(s), \boldsymbol{\mu}'(s), \boldsymbol{\lambda}, \boldsymbol{\phi}(s), \xi(s), \boldsymbol{\zeta}(s)\}$. The optimal policy function is then a function $z(y)$ that satisfies the relationship $F(z(y), y, g(z(y))) = 0$. Taking total derivatives around the steady state $\overline{y}$ and $\overline{z} = z(\overline{y})$

$$D_z F(\overline{z}, \overline{y}, g(\overline{z}))D_y z(\overline{y}) + D_y F(\overline{z}, \overline{y}, g(\overline{z})) + D_g F(\overline{z}, \overline{y}, g(\overline{z}))Dg(\overline{z})D_y z(\overline{z}) = 0$$

In order to linearize $z(y)$ around the steady state $\overline{y}$ we need to compute $D_y z(\overline{y})$. The envelope condition (77b) tell us that $V_{\boldsymbol{\rho}}$ can be computed from the optimal policies, i.e.

$$\begin{pmatrix} \boldsymbol{x}(\boldsymbol{\mu}, \boldsymbol{\rho}) \\ V_{\boldsymbol{\rho}}(\boldsymbol{\mu}, \boldsymbol{\rho}) \end{pmatrix} = w(z(\boldsymbol{\mu}, \boldsymbol{\rho})) = \begin{pmatrix} \boldsymbol{x} \\ -\boldsymbol{\lambda}\mathbb{E}\left[U_{c,1}\right] \end{pmatrix}$$

If we let Φ_s be the matrix that maps $z(\boldsymbol{\mu}, \boldsymbol{\rho})$ into $\begin{pmatrix} \boldsymbol{\mu}'(s) \\ \boldsymbol{\rho}'(s) \end{pmatrix}$ then we can write $g(\boldsymbol{\mu}, \rho)$ using z and w as follows

$$g(z) = \begin{pmatrix} w(z(\Phi_1 z)) \\ w(z(\Phi_2 z)) \end{pmatrix}$$

taking derivatives we quickly obtain that

$$\begin{aligned} D_z g(\overline{z}) &= \begin{pmatrix} Dw(z(\Phi_1\overline{z})) & 0 \\ 0 & Dw(z(\Phi_2\overline{z})) \end{pmatrix} \begin{pmatrix} D_y z(\Phi_1\overline{z}) & 0 \\ 0 & D_y z(\Phi_1\overline{z}) \end{pmatrix} \underbrace{\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}}_{\Phi} \\ &= \begin{pmatrix} Dw(\overline{z}) & 0 \\ 0 & Dw(\overline{z}) \end{pmatrix} \begin{pmatrix} D_y z(\overline{y}) & 0 \\ 0 & D_y z(\overline{y}) \end{pmatrix} \Phi \\ &= \begin{pmatrix} Dw(\overline{z})D_y z(\overline{y}) & 0 \\ 0 & Dw(\overline{z})D_y z(\overline{y}) \end{pmatrix} \Phi \end{aligned}$$

We can then go back to our original matrix equation to obtain

$$D_z F(\overline{z}, \overline{y}, \overline{w}) D_y z(\overline{y}) + D_y F(\overline{z}, \overline{y}, \overline{w}) + D_w F(\overline{z}, \overline{y}, \overline{w}) \begin{pmatrix} Dw(\overline{z}) D_y z(\overline{y}) & 0 \\ 0 & Dw(\overline{z}) D_y z(\overline{y}) \end{pmatrix} \Phi D_y z(\overline{z}) = 0 \quad (79)$$

, where $\overline{w} = g(\overline{z}) = w(\overline{z})$. This is now a non-linear matrix equation for $D_y z(\overline{y})$, where all the other terms can be computed using the steady state values $\overline{z}$ and $\overline{y}$ (note $g(\overline{z})$ is known from the envelope conditions at the steady state). Furthermore, $D_y z(\overline{y})$ gives us the linearization of the policy rules since to first order

$$z \approx \overline{z} + D_y z(\overline{y})(y - \overline{y})$$

Our procedure for computing the linearization proceeds as follows

1. Find the steady state by solving the system of equations (26. Numerically, we have found that this is very robust to the parameters of the model.

2. Compute $D_z F(\overline{z}, \overline{y}, g(\overline{z}))$, $D_z F(\overline{z}, \overline{y}, g(\overline{z}))$ and $D_v F(\overline{z}, \overline{y}, g(\overline{z}))$ by numerically differentiating F. This is straightforward using auto-differentiation.

3. Compute $Dw(\overline{z})$ using auto-differentiation.

4. Construct a matrix equation as follows. Given policies $A = Dw(\overline{z}) D_y z(\overline{y})$ (these are the linearized policies of $\boldsymbol{x}$ and $V_{\boldsymbol{\rho}}$ with respect to $(\boldsymbol{\mu}, \boldsymbol{\rho})$), it is possible to solve for $D_y z(\overline{y})$ from

$$D_y z(\overline{z}) = -\left(D_z F(\overline{z}, \overline{y}, \overline{w}) + D_w F(\overline{z}, \overline{y}, \overline{w}) \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \Phi \right)^{-1} D_y F(\overline{z}, \overline{y}, \overline{w})$$

   We wish to find an A such that

$$A = Dw(\overline{z}) D_y z(\overline{z})$$

Given the linearized policy rules it is then possible to evaluate the local stability of the steady state. We find that in the absence of discount factor shocks the steady state is stable generically across the parameter space.

This linearization can be used to construct the bordered hessian of the problem (23) at the steady state. We can then apply second order tests to verify that the the first order necessary conditions are sufficient.

9.6 Proof for Proposition 6

Proof.

The state at time t can be written as

$$\hat{\Psi}_t = B_t B_{t-1} \cdots B_1 * \hat{\Psi}_0.$$

where the B_i are all random variables being $B(s)$ with probability $\Pr(s)$. Taking expectations and applying independence we then obtain

$$\mathbb{E}_0[\hat{\Psi}_t] = \mathbb{E}_0[B_t B_{t-1} \cdots B_1] \hat{\Psi}_0 \quad (80)$$

$$\mathbb{E}[B_t] \mathbb{E}[B_{t-1}] \cdot \mathbb{E}[B_1] \hat{\Psi}_0 \quad (81)$$

$$\overline{B}^t \hat{\Psi}_0 \quad (82)$$

where $\overline{B} = \mathbb{E}B(s)$. If eigenvalues of $\overline{B}$ are positive and strictly less than 1, at least, in expectation the linearized system converges that is

$$\bar{\hat{\Psi}}_t = \mathbb{E}_0[\hat{\Psi}_t] = \overline{B}^t \hat{\Psi}_0 \to \mathbf{0}. \tag{83}$$

It should be noted that the conditional expectation actually captures a significant portion of the linearized dynamics. The remaining question is does the distribution converge to $\mathbf{0}$. This can be done by analyzing the variance. Let

$$\Sigma_{\Psi,t} = \mathbb{E}_0 \left[(\hat{\Psi}_t - \bar{\hat{\Psi}}_t)(\hat{\Psi}_t - \bar{\hat{\Psi}}_t)' \right]$$

or

$$\Sigma_{\Psi,t} = \mathbb{E}_0 \hat{\Psi}_t \hat{\Psi}_t' - \bar{\hat{\Psi}}_t \bar{\hat{\Psi}}_t'. \tag{84}$$

Note that if eigenvalues of $\overline{B}$ are positive and strictly less that 1, $\bar{\hat{\Psi}}_t$ converges to 0. Using the independence of $\hat{\Psi}_{t-1}$ and B_t, and the identities $\hat{\Psi}_t = B_t \hat{\Psi}_{t-1}$ and $\bar{\hat{\Psi}}_t = \overline{B}\Psi_{t-1}$ we quickly obtain that for large t

$$\Sigma_{\Psi,t} \approx \mathbb{E}[B \Sigma_{\Psi,t-1} B'.] \tag{85}$$

Showing that $\hat{\Psi}_t \to \mathbf{0}$ in probability, amounts to showing that $\Sigma_{\Psi,t} \to 0$ for any starting point Σ_Ψ and following the process in equation (85). One can obtain a necessary condition for $\|\Sigma_{\Psi,t}\| \to 0$ under the process in equation (85). That process can be rewritten as follows

$$\begin{aligned}
\Sigma_{\Psi,t} &= \mathbb{E}[B \Sigma_{\Psi,t-1} B'] && (86)\\
&= \sum_s \Pr(s) B(s) \Sigma_{\Psi,t-1} B(s)' && (87)\\
&= \sum_s \Pr(s)(\overline{B} + (B(s) - \overline{B}))\Sigma_{\Psi,t-1}(\overline{B} + (B(s) - \overline{B}))' && (88)\\
&= \overline{B}\Sigma_{\Psi,t-1}\overline{B}' + \sum_s \Pr(s)(B(s) - \overline{B})\Sigma_{\Psi,t-1}(B(s) - \overline{B})'. && (89)
\end{aligned}$$

This is a deterministic linear system in $\Sigma_{\Psi,t}$. Suppose we reshape $\Sigma_{\Psi,t}$ as a vector (denoted by $\text{vec}(\Sigma_{\Psi,t})$) and let $\hat{B}$ be a (square) matrix such that equation 89 is written as

$$\text{vec}(\Sigma_{\Psi,t}) = \hat{B}\text{vec}(\Sigma_{\Psi,t-1}).$$

The stability of this system is guaranteed if the (real part) of eigenvalues of $\hat{B}$ are less than 1. ■

References

Aiyagari, S. Rao, and Ellen R. McGrattan. 1998. "The optimum quantity of debt." *Journal of Monetary Economics*, 42(May 1997): 447–469, URL: http://www.sciencedirect.com/science/article/pii/S0304393298000312.

Albanesi, S., and R. Armenter. 2012. "Intertemporal Distortions in the Second Best." *The Review of Economic Studies*, 79(4): 1271–1307, URL: http://restud.oxfordjournals.org/lookup/doi/10.1093/restud/rds014, DOI: http://dx.doi.org/10.1093/restud/rds014.

Autor, DH, LF Katz, and MS Kearney. 2008. "Trends in US wage inequality: Revising the revisionists." *The Review of Economics and Statistics*, 90(2): 300–323, URL: http://www.mitpressjournals.org/doi/abs/10.1162/rest.90.2.300, DOI: http://dx.doi.org/10.1162/rest.90.2.300.

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.

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.

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.

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.

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

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.

Battaglini, Marco, and Stephen Coate. 2007. "Inefficiency in Legislative Policymaking: A Dynamic Analysis." *American Economic Review*, 97(1): 118–149, URL: http://www.aeaweb.org/articles.php?doi=10.1257/aer.97.1.118, DOI: http://dx.doi.org/10.1257/aer.97.1.118.

Battaglini, Marco, and Stephen Coate. 2008. "A Dynamic Theory of Public Spending, Taxation, and Debt." *American Economic Review*, 98(1): 201–236, URL: http://www.aeaweb.org/articles.php?doi=10.1257/aer.98.1.201, DOI: http://dx.doi.org/10.1257/aer.98.1.201.

Boldrin, Michele, Lawrence J. Christiano, and Jonas DM Fisher. 2001. "Habit Persistence, Asset Returns, and the Business Cycle." *American Economic Review*, 91(1): 149–166, URL: `http://ideas.repec.org/a/aea/aecrev/v91y2001i1p149-166.html`.

Bryant, John, and Neil Wallace. 1984. "A Price Discrimination Analysis of Monetary Policy." *The Review of Economic Studies*, 51(2): , p. 279, URL: `http://restud.oxfordjournals.org/lookup/doi/10.2307/2297692`, DOI: `http://dx.doi.org/10.2307/2297692`.

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`.

Constantinides, George, and Darrell Duffie. 1996. "Asset pricing with heterogeneous consumers." *Journal of Political economy*, 104(2): 219–240, URL: `http://www.jstor.org/stable/10.2307/2138925`.

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`.

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

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`.

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`.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2012. "The Nature of Countercyclical Income Risk." Working Paper 18035, National Bureau of Economic Research.

Heathcote, Jonathan. 2005. "Fiscal Policy with Heterogeneous Agents and Incomplete Markets." *Review of Economic Studies*, 72(1): 161–188, URL: `http://restud.oxfordjournals.org/lookup/doi/10.1111/0034-6527.00328`, DOI: `http://dx.doi.org/10.1111/0034-6527.00328`.

Judd, Kenneth L., Lilia Maliar, and Serguei Maliar. 2011. "Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models." *Quantitative Economics*, 2(2): 173–210, URL: `http://doi.wiley.com/10.3982/QE14`, DOI: `http://dx.doi.org/10.3982/QE14`.

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.

Lucas Jr., 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.

Magill, Michael, and Martine Quinzii. 1994. "Infinite Horizon Incomplete Markets." *Econometrica*, 62(4): 853–880, URL: http://www.jstor.org/stable/2951735, DOI: http://dx.doi.org/10.2307/2951735.

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

Ray, Debraj. 2002. "The Time Structure of Self-Enforcing Agreements." *Econometrica*, 70(2): 547–582, URL: http://onlinelibrary.wiley.com/doi/10.1111/1468-0262.00295/full.

Reinhart, Carmen M, and Kenneth S Rogoff. 2010. "Growth in a Time of Debt." *American Economic Review*, 100(2): 573–578, URL: http://pubs.aeaweb.org/doi/abs/10.1257/aer.100.2.573, DOI: http://dx.doi.org/10.1257/aer.100.2.573.

Sargent, Thomas J. 1987. *Dynamic macroeconomic theory.*: Harvard University Press.

Sargent, Thomas J., and Bruce D. Smith. 1987. "Irrelevance of open market operations in some economies with government currency being dominated in rate of return." *American Economic Review*, 77(1): 78–92, URL: http://ideas.repec.org/a/aea/aecrev/v77y1987i1p78-92.html.

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

Wallace, Neil. 1981. "A Modigliani-Miller Theorem for Open-Market Operations." *The American Economic Review*, 71(3): 267–274, URL: http://www.jstor.org/stable/1802777, DOI: http://dx.doi.org/10.2307/1802777.

Werning, Iván. 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.

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

Yared, Pierre. 2013. "Public Debt Under Limited Private Credit." *Journal of the European Economic Association*, 11(2): 229–245, URL: http://doi.wiley.com/10.1111/jeea.12010, DOI: http://dx.doi.org/10.1111/jeea.12010.