Taxes, debts, and redistributions with aggregate shocks*

Anmol Bhandari
bhandari@umn.edu

David Evans
dgevans@nyu.edu

Mikhail Golosov
golosov@princeton.edu

Thomas J. Sargent
thomas.sargent@nyu.edu

November 2014

Abstract

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

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

*We thank Mark Aguiar, Stefania Albanesi, Manuel Amador, Andrew Atkeson, Marco Bassetto, V.V. Chari, Harold L. Cole, Guy Laroque, Francesco Lippi, 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

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

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

Section 2 sets out preferences and possibilities that define the economic environment. Section 3 describes a Ricardian property for our environment that we subsequently exploit in concisely formulating the Ramsey problem in section 4. Net, not gross asset positions, not affect the set of allocations that can be implemented in competitive equilibria with a proportional labor tax and transfers. This insight reduces the dimension of the state needed to characterize a Ramsey plan recursively. It also justifies a normalization that lets us interpret transfers and public debt separately.

A Ramsey plan induces an ergodic distribution of transfers, the labor tax rate, and the distribution of assets that is determined by interactions between a) the government's ability to hedge aggregate government expenditure shocks by taking advantage of fluctuations in the return on the single asset; and b) the government's preferences about redistribution. The analysis in sections 5, 6, and 7 shows that these interactions shape the ergodic distribution of government debt in the following ways. If equilibrium outcomes make the return on the asset *low* when the net-of interest government deficit is high, then the government will run up debt. But if the return on the asset is *high* when the net-of-interest government deficit is high, the government will accumulate assets. The long run variances of government assets and the tax rate are both lower and rates of convergence to the ergodic distribution are higher in economies in which the comovement between the net-of interest deficit and the return on the asset is bigger in absolute magnitude. Other things equal, governments that want more redistribution eventually issue more debt.

The comovement between the aggregate shock that drives government expenditures and returns on the single asset is a key intermediating object that shapes the ergodic joint distribution of the tax rate and government debt. In our general setting, one-period utilities are concave in consumption. That makes the return on the single asset an e equilibrium objects partly under the control of the Ramsey planner. To help us understand forces shaping both the ergodic joint distribution and rates of convergence to it, we begin by studying a special setting in which returns on the single asset and their correlation with the aggregate shock are exogenous, namely, a setting with one-period utilities that are quasilinear, meaning linear in consumption. Section 5 executes this quasilinear analysis in two parts. Subsection 5.1 begins by analyzing a representative agent economy with quasilinear preferences, a government restricted to set transfers equal to zero always, and i.i.d aggregate shocks. Our main finding here is that for a large class of payoff structures, debt drifts towards an ergodic set that in a sense maximizes the government's ability to hedge fiscal shocks, a set that primarily depends on how the payoff on the asset correlate with fluctuations in the government's net-of-interest deficit. In particular, if the payoff is high when the government needs revenue to finance a higher net-of-interest deficit, optimal hedging requires the government to issue positive debt, and, conversely, if the asset payoff is low in such times, the government would want to hold positive assets. The magnitude of debt (or assets) is decreasing and the speed at which the debt converges is increasing in the magnitude of this comovement. For special cases in which the asset payoff is affine in government expenditure shocks, we show that the ergodic distribution is degenerate. For other assumptions about the asset payoff, shock correlation, we develop tools to approximate the ergodic distribution and tell how the spread of the ergodic distribution of government debt and the tax rate increases with how far the payoffs are from allowing perfect fiscal hedging.

The section 5.1 representative agent analysis with transfers restricted to zero turns out to be informative about what is going on in multiple agent economies with no restrictions on transfers but with Pareto weights that make the welfare costs of transfers so high that the Ramsey planner chooses never to use them, or at least not to use them eventually.[1] Subsection 5.2 establishes this conclusion by analyzing an economy with unrestricted transfers, and two types of agents both of whom have quasilinear one-period utilities and one of whom is not productive. We study the effects of the presence or absence of a nonnegativity constraint on the consumption of the unproductive agent as a determinant of the asymptotic level of government debt and whether and when transfers are used. The asymptotic level of assets is decreasing in the planner's desire for redistribution. This comes from the fact that welfare costs of using transfers are lower for a more redistributive government. Consequently it relies more on transfers and has less cause to accumulate assets to hedge aggregate shocks.

In section 6, we study economies that are more general in terms of their heterogeneity, preferences, and shock structures. Subsection 6.1 reformulates the section 4 Ramsey problem recursively in terms of two Bellman equation, one for time 0 and another for times $t \geq 1$. That these Bellman equations have different state variables expresses the time inconsistency of our Ramsey plans, which as usual attribute to the time 0 Ramsey planner the ability to commit to an intertemporal plan. Subsections 6.2 and 6.3 study a special case with binary iid shocks and characterizes conditions under which the planner eventually achieves complete hedging by having the appropriate constant level of government debt or assets while managing to keep both the tax rate on labor and transfers constant. Subsection 6.4 analyzes a more general set of economies with iid binary shocks **Anmol XXXXX: let's fill in here the sense in which it is more general again – e.g., shock process or utilities or number of types of agents.** by applying an extension of the same type of approximation method used earlier in section 5.1. also used in Theorem 4 approximation method.

In section 7, we numerically verify that the forces isolated in the more analytically tractable models of sections 5 and 6 extend to a version of the model with several types of agents calibrated to match US data. These insights extend to economies with preferences having curvature in the utilities from consumption. We calibrate a version of our model to US data that captures (1) the initial heterogeneity wages and assets; (2) the observation that in recessions the left tail of the cross-section distribution of labor income falls by more than right tail; and (3) how inflation and asset return risk comove with labor productivity. We use this model to validate and quantify the importance of the different channels that were emphasized in our theoretical analysis with simpler environments. Besides this we also describe features of optimal government policy, especially in booms and recessions at higher frequencies. We find that during recessions

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

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

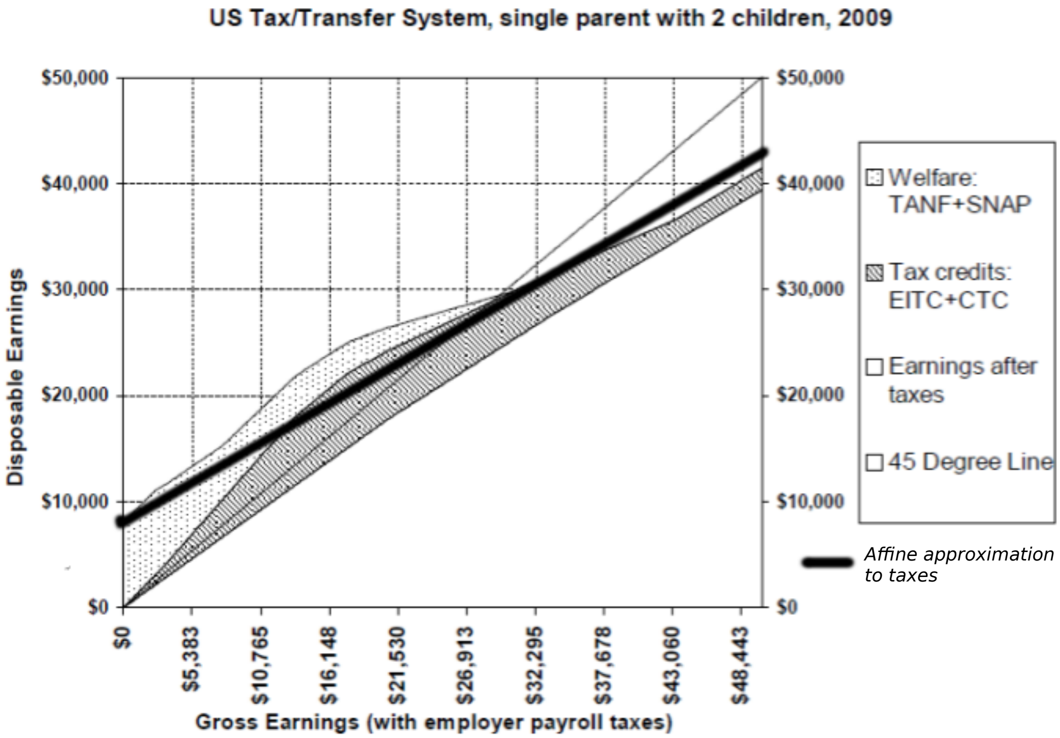
In this paper, we have adhered to the adage that technique ought to efface itself in the service of economic analysis. Nevertheless, it is appropriate to mention that we have attained the main results in this paper only by taking advantage of some important technical improvements developed during our research for this paper. Two types contributions are interesting in their own right because they will surely be useful in other applications. They are (1) a set of efficient computer programs for approximating the section 6.1 Bellman equations by using projection methods to form XXXXXX approximations of the XXXXXXX; **Anmol: XXXXX let's fill in precisely what objects were calculated here – e.g., the multipliers as functions of XXXX using XXXXX polynomials.** ; and (b) distinct applications in sections 6.4, 4, and 7 of a new approach of Evans (2014) to computing linear approximations of dynamic incomplete markets economies. Where possible, we have used methods of type 1 to estimate and confirm the accuracy of the type 2 approximations. Quantitative applications as ambitious as those in our section 7 are made possible by our use of the type 2 method. An economic insight underlies the type 2 method – that along a sample path of an incomplete markets economy, there are economic reasons to recommend taking a sequence linear or polynomial approximations about the steady states of a sequence of *other* economies and evaluating the polynomial at the current wealth distribution of the incomplete markets economy. This insight underlying the type 2 method did not appear out of thin air but instead emerged partly from studying the structure of the economic model in this paper.

Section 8 offers concluding remarks. The paper ends with 55 appendixes.

1.1 Relationships to literatures

Anmol and Mike: this subsection has just been cut and pasted from a long ago version. I don't like literature review sections, but I think we need one for strategic reasons if only to help an editor get us fair referees.

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



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

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

Several other papers impute distributive concerns to a Ramsey planner. Three papers most

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

closely related to ours are Bassetto (1999), Shin (2006), and Werning (2007). Like us, those authors allow heterogeneity and study distributional consequences of alternative tax and borrowing policies. Bassetto (1999) extends the Lucas and Stokey (1983) environment to include N types of agents with heterogeneous time-invariant labor productivities. There are complete markets. The Ramsey planner has access only to proportional taxes on labor income and state-contingent borrowing. Bassetto studies how the Ramsey planner's vector of Pareto weights influences how he responds to government expenditures and other shocks by adjusting the proportional labor tax and government borrowing to cover expenses while manipulating competitive equilibrium prices to redistribute wealth between 'rentiers' (who have low productivities and whose main income is from their asset holdings) and 'workers' (who have high productivities) whose main income source is their labor.

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

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

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

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

Finally, our numerical analysis in Section 7 is related to McKay and Reis (2013). While our focus differs from theirs – McKay and Reis study the effect of a calibrated version of the US tax and transfer system on stabilization of output, while we focus on optimal policy in a simpler economy – both papers confirm the importance of transfers and redistribution over business-cycle frequencies.

2 Environment

Exogenous fundamentals include a stochastic cross section of skills $\{\theta_{i,t}\}$, government expenditures g_t, and the payoff p_t on an asset. These are all functions of a shock $s_t \in S$ governed by an irreducible Markov process; S is a finite set. We let $s^t = (s_0, ..., s_t)$ denote a history of shocks having joint probability density $\pi(s^t)$.[5]

There is a mass n_i of type $i \in I$ agents, with $\sum_{i=1}^{I} n_i = 1$. Types differ in skills indexed by $\{\theta_{i,t}\}_t$. 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

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

where $\mathbb{E}_t$ is a mathematical expectations operator conditioned on time t information and $\beta \in (0, 1)$ is a time discount factor. Except in sections 3 and 5, we assume that $U^i : \mathbb{R}^2_+ \to \mathbb{R}$ is concave in $(c, -l)$ and twice continuously differentiable. We let $U^i_{x,t}$ or $U^i_{xy,t}$ denote first and second derivatives of U^i with respect to $x, y \in \{c, l\}$ in period t and assume that for all (c, i) the $\lim_{x\to 0} U^i_l (c, x) = 0$. The Ricardian flavored results in section 3 hold under weaker assumptions about U^i, while section 5 assumes quasi-linear preferences in order to isolate key forces analytically.

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

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

Households and the government trade a single, possibly risky, asset whose time t payoff p_t is described by

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

[5]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.

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

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

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

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

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

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

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

and the government budget constraint is

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

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

Remark 1 *We impose debt limits on asset profiles.*[6]

Definition 2 *For a given initial asset distribution* $\left(\{b_{i,-1}\}_i, B_{-1}\right)$*, a competitive equilibrium with affine taxes is a sequence* $\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, B_t, R_t\}_t$ *and a tax policy* $\{\tau_t, T_t\}_t$ *such that (i)*

[6]For households, we shall impose natural debt limits that depend on the tax policy. An alternative is to impose ad-hoc debt limits in the form of exogenous history-contingent bounds for each agent. Appendix A.1 discusses how restricting attention to natural debt limits for the households shrinks the set of allocations that can be implemented as competitive equilibria.

$\{c_{i,t}, l_{i,t}, b_{i,t}\}_{i,t}$ *maximize (1) subject to (4) and the condition that* $\{b_{i,t}\}_{i,t}$ *satisfies the borrowing limits; and (ii) constraints (2), (3),and (5) are satisfied.*

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

$$\mathbb{E}_0 \sum_{i=1}^{I} \omega_i \sum_{t=0}^{\infty} \beta^t U_t^i \left(c_{i,t}, l_{i,t}\right), \tag{6}$$

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

Definition 3 *Given* $(\{b_{i,-1}\}_i, B_{-1})$*, an optimal competitive equilibrium with affine taxes is a competitive equilibrium with an allocation that maximizes* (6)*.*

3 Ricardian equivalence

The arithmetic of budget constraints and market clearing instructs us how to formulate the optimal policy problem concisely. An equivalence class of tax policies and asset profiles supports the same competitive equilibrium allocation.

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

We relegate the proof to appendix A.2. In the spirit of Barro (1974), we interpret Theorem 1 as asserting a type of Ricardian equivalence. Similar results hold in more general environments. For example, we could allow agents to trade all conceivable Arrow securities and still show that equilibrium allocations depend only on agents' net assets positions.

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*

$$b'_{i,-1} - b'_{1,-1} = b''_{i,-1} - b''_{1,-1} \ \forall i.$$

Thus, total government debt is not what matters, who owns it does. To illustrate this point, imagine an 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 losses are convex in τ, higher levels of debt financed with bigger distorting taxes $\{\tau_t\}$ impose larger distortions and thereby degrade 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 suitable transfers, we need to know who owns the initial government debt. For example, suppose that agents own equal amounts. 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 of transfers by exactly the amount of the increase in per capita debt. Then the labor tax rate sequence $\{\tau_t\}$ and the allocation could both remain unchanged.[7]

The situation would be different if holdings of government debt were not equal across agents. For example, suppose that richer people initially own disproportionately more government debt. 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 $\{\omega_i\}$ that favor equality would want to increase both transfers $\{T_t\}$ and the distorting labor tax rate $\{\tau_t\}$ 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 confirms the importance of the distribution of government debt across people. Government debt that is widely and evenly distributed (e.g., implicit Social Security debt) is less distorting than government debt owned mostly by people whose incomes are at the top of the income distribution (e.g., government debt held by hedge funds).[8]

Throughout this paper we avail ourselves of theorem 1 to impose a normalization on asset profiles: we assume that productivities are ordered as $\theta_{1,t} \geq \theta_{2,t} \ldots \geq \theta_{I,t}$ and set $b_{I,t} = 0$. This allows us to interpret $-B_t = \sum_{i<I} n_i b_{i,t}$ as public debt. Limits on $\sum_{i<I} n_i b_{i,t}$ are counterparts to debt limits in a representative agent economy.

[7] This example illustrates principles proclaimed by Simon Newcomb (1865, p. 85) in the quotation that begins this paper.

[8] It is possible to extend our analysis to open economy with foreign holdings of domestic debt. The more government debt is owned by the foreigners, the higher are the distorting taxes that the government needs to impose.

4 Optimal equilibria with affine taxes

With natural borrowing limits for households, first-order necessary conditions for the household's problem are

$$(1-\tau_t)\,\theta_{i,t}U^i_{c,t} = -U^i_{l,t}, \tag{7}$$

and

$$U^i_{c,t} = \beta\mathbb{E}_t R_{t+1}U^i_{c,t+1}. \tag{8}$$

To help characterize an equilibrium, we use

Theorem 2 *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), (7), and (8) and $b_{i,t}$ is bounded for all i and t.*

Proof. Necessity is obvious. In appendix A.3, 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), (7), and (8) solves consumer i's problem. Equilibrium $\{B_t\}_t$ is determined by (3) and constraint (5) is then implied by Walras' Law ■

By Theorem 2, to find an optimal equilibrium we can choose $\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, R_t, \tau_t, T_t\}_t$ to maximize (6) subject to (2), (4), (7), and (8). We apply a first-order approach and follow steps similar to those taken by Lucas and Stokey (1983) and Aiyagari et al. (2002). Substituting consumers' first-order conditions (7) and (8) into the budget constraints (4) yields the implementability constraints

$$c_{i,t} + b_{i,t} = -\frac{U^i_{l,t}}{U^i_{c,t}} l_{i,t} + T_t + \frac{p_t U^i_{c,t-1}}{\beta\mathbb{E}_{t-1}p_t U^i_{c,t}} b_{i,t-1} \;\forall i \geq 1, t \geq 1, \tag{9}$$

and

$$c_{i,0} + b_{i,0} = -\frac{U^i_{l,0}}{U^i_{c,0}} l_{i,0} + T_0 + p_0\beta^{-1} b_{i,-1} \;\forall i \geq 1. \tag{10}$$

For $I \geq 2$, we can use constraints (9) and (10) for $i = 1$ to eliminate T_t for $i < I$. Letting $\tilde{b}_{i,t} \equiv b_{i,t} - b_{I,t}$, we can represent the implementability constraints as

$$\begin{aligned} &(c_{i,t} - c_{I,t}) + \tilde{b}_{i,t} \\ = \;& -\frac{U^i_{l,t}}{U^i_{c,t}} l_{i,t} + \frac{U^1_{l,t}}{U^1_{c,t}} l_{1,t} + \frac{p_t U^i_{c,t-1}}{\beta\mathbb{E}_{t-1}p_t U^i_{c,t}} \tilde{b}_{i,t-1} \text{ for } \; i > 1 \text{ and } \; t \geq 1, \end{aligned} \tag{11}$$

so that the planner's maximization problem involves only on the $I-1$ variables $\tilde{b}_{i,t-1}$.

Denote $Z_t^i = (c_{i,t} - c_{I,t}) + \tilde{b}_{i,t} + \frac{U^i_{l,t}}{U^i_{c,t}} l_{i,t} - \frac{U^I_{l,t}}{U^I_{c,t}} l_{1,t}$. The Ramsey problem is:[9]

$$\max_{\{c_{i,t}, l_{i,t}, \tilde{b}_{i,t}\}} \mathbb{E}_0 \sum_{i=1}^{I} \omega_i \sum_{t=0}^{\infty} \beta^t U^i \left(c_{i,t}, l_{i,t}\right) \tag{12}$$

subject to

$$\tilde{b}_{i,t-1} \frac{p_t U^i_{c,t-1}}{\mathbb{E}_{t-1} p_t U^i_{c,t}} = \mathbb{E}_t \sum_{k=t}^{\infty} \beta^{k-t} \left(\frac{U^i_{c,k}}{U^i_{c,t}}\right) Z^i_k \quad \forall t \geq 1, i < I \tag{13a}$$

$$\tilde{b}_{i,-1} p_0 = \mathbb{E}_{-1} \sum_{k=0}^{\infty} \beta^k \left(\frac{U^i_{c,k}}{U^i_{c,t}}\right) Z^i_k \quad \forall i < I \tag{13b}$$

$$\frac{\mathbb{E}_t p_{t+1} U^i_{c,t+1}}{U^i_{c,t}} = \frac{\mathbb{E}_t p_{t+1} U^I_{c,t+1}}{U^I_{c,t}} \quad \forall t \geq 1, i < I \tag{13c}$$

$$\sum_{i=1}^{I} n_i c_i(s^t) + g\left(s_t\right) = \sum_{i=1}^{I} n_i \theta_i \left(s_t\right) l_i(s^t) \quad \forall t \geq 0 \tag{13d}$$

$$\frac{U^i_{l,t}}{\theta_{i,t} U^i_{c,t}} = \frac{U^I_{l,t}}{\theta_{I,t} U^I_{c,t}} \quad \forall t \geq 0, i < I \tag{13e}$$

$$\sum_{i<I} \tilde{b}_{i,t-1} \text{ is bounded} \quad \forall t \geq 0 \tag{13f}$$

In section 5, we exploit some simplifications that come with quasilinear utility, but we return to this more general formulation in sections 6 and 7.

5 Quasilinear preferences

We specialize the section 4 Ramsey problem by maintaining the following assumptions throughout this section.[10]

Assumption 1 *IID shocks to expenditure: $g(s_t)$ is i.i.d over time*

Assumption 2 *Quasilinear preferences: $u(c,l) = c - \frac{l^{1+\gamma}}{1+\gamma}$*

With i.i.d shocks we can restrict our attention to payoff matrices $\mathbb{P}$ that have identical rows denoted by a vector $P(s)$ with a normalization $\mathbb{E}P(s) = 1$.

The assumption of quasilinear utilities effectively makes the return on the single asset exogenous. That allows us to focus directly on how exogenous comovements in the asset return with

[9] When $p_t \equiv 1$, so that the asset is risk-free, constraint (13a) requires that the conditional expectation at time t on the right side be an exact function of information at time $t-1$. This is the measurability condition that lies at the heart of Aiyagari et al. (2002). More generally, condition (13a) is inherited from the restriction that only one asset is traded and that it has payoff p_t.

[10] Aiyagari et al. (2002) assume quasilinear preferences in an important part of their analysis.

the aggregate shock impinge on the Ramsey planner's incentives to issue debt or accumulate assets. One way to read some of the results in this section is that in various ways they extend Aiyagari et al.'s (2002) analysis of a representative agent economy with quasilinear preferences in which the government is allowed to use only nonnegative transfers. Before studying a Ramsey allocation for an economy with heterogeneous agents and no restrictions on transfers in subsection 5.2, we find it instructive in subsection 5.1 to study a representative agent economy in which the government *cannot* use transfers. This representative agent economy is informative about multiple-agent economies allowing transfers but with Pareto weights that imply that the welfare costs of transfers are too high ever to use. The results also shed light on how market incompleteness, as captured by the structure of $P(s)$, impedes tax-smoothing in a joint ergodic distribution for the tax rate, transfers, and government debt.

5.1 Representative agent

We alter the Aiyagari et al. (2002) economy in two ways: first, the single asset can be risky; second, government transfers are constrained to be zero always. Both features contribute to modifying the Aiyagari et al. result that in the long run the tax rate on labor is zero.[11]

Given a tax, asset policy $\{\tau_t, B_t\}$, the household solves,

$$\max_{\{c_t, l_t, b_t\}_t} \mathbb{E}_0 \sum_t \beta^t \left[c_t - \frac{l_t^{1+\gamma}}{1+\gamma} \right] \tag{14}$$

subject to

$$c_t + b_t = (1-\tau_t)\theta l_t + R_t b_{t-1}, \quad t \geq 0 \tag{15}$$

and the debt limits $-\overline{B} \geq b_t \leq -\underline{B}$.[12] At interior solutions, the first-order conditions for l and b imply that

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

where we are again using the normalization $E_{t-1} p_t = 1$. Using these equations to express the tax rate and asset return as functions c and l and then substituting them into the budget constraint (15) gives

$$c_t + b_t = l_t^{1+\gamma} + \frac{p_t}{\beta} b_{t-1}, \quad t \geq 0. \tag{16}$$

[11] Aiyagari et al. allow nonnegative transfers.

[12] We remind the reader of the convention that we set $R_0 = \beta^{-1} P(s_0)$.

Equation (16) presents a recursive version of the implementability constraints (13a) appropriate for the present setting. We also have the feasibility constraints

$$c_t + g_t \leq \theta l_t, \ \forall t \geq 0 \tag{17a}$$

and the market-clearing conditions for bonds,

$$b_t + B_t = 0, \ \forall t \geq 0. \tag{17b}$$

Given $b_{-1} = -B_{-1}$, the Ramsey planner chooses an allocation and a government debt sequence $\{B_t\}_{t=0}^{\infty}$ subject to the implementability conditions (16), feasibility (17a), market clearing for bonds (17b), and limits $(\underline{B}, \overline{B})$ on government assets.[13]

Let $V(B_{\_})$ be the ex-ante value of a Ramsey plan starting with initial government assets $B_{\_}$. It satisfies the Bellman equation

$$V(B_{\_}) = \max_{\{c(s),l(s),B(s)\}_s} \sum_s \pi(s) \left\{ c(s) - \frac{l(s)^{1+\gamma}}{1+\gamma} + \beta V(B(s)) \right\} \tag{18}$$

where the maximization is subject to

$$c(s) - B(s) = l(s)^{1+\gamma} - \beta^{-1} P(s) B_{\_} \tag{19a}$$

$$c(s) + g(s) \leq \theta l(s) \tag{19b}$$

$$\underline{B} \leq B(s) \leq \overline{B} \tag{19c}$$

where $\overline{B}$ and $\underline{B}$ are upper and lower bounds for government assets.

To organize analytical results, it is useful to collect some $P(s)$ vectors that are perfectly correlated with expenditure shocks $g(s)$ in the following set indexed by a constant B^*

$$\mathcal{P}^* = \left\{ P(s) : P(s) = 1 + \frac{\beta}{B^*}(g(s) - \mathbb{E}g) \text{ for some } B^* \in [\overline{B}, \underline{B}] \right\}. \tag{20}$$

Theorem 3 *In a representative agent economy satisfying assumptions 1 and 2, the behavior of government assets under a Ramsey plan can be characterized as follows:*

1. *Suppose $P(s) \notin \mathcal{P}^*$. There is an invariant distribution of government assets such that*

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

[13] In some calculations, we will impose a natural debt limit $\underline{B}$ for the government.

2. *Suppose* $P(s) - P(s') > \beta \frac{g(s)-g(s')}{-\underline{B}} \quad \forall s, s'$. *The value function* $V(B_{\_})$ *is strictly concave and there exist* $B_1 < B_2$ *such that*

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

and

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

These inequalities imply that for large enough government assets (or debt), assets drift towards an interior region.

3. *Suppose* $P(s) \in \mathcal{P}^*$. *Government assets converge to a degenerate steady state*

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

where B^* *is the object appearing in definition* (20) *of* $\mathcal{P}^*$ *and satisfies,*

$$B^* = \beta \frac{\text{var}(g(s))}{\text{cov}(P(s), g(s))}. \tag{21}$$

The long-run tax rate is inversely related to B^* *and satisfies:*

$$\lim_{B^* \to \underline{B}} \tau^* = \frac{\gamma}{1+\gamma}, \quad \lim_{B^* \to \infty} \tau^* = -\infty.$$

Theorem 3 describes how the invariant distribution of government assets depends on the payoff vector $P(s)$. Part 1 of theorem 3 establishes that when the payoff vector does not belong to $\mathcal{P}^*$, the support of the invariant distribution of assets is wide in the sense that almost all asset sequences recurrently revisit small neighborhoods of any arbitrary lower and upper bounds on government assets. Because the labor tax rate is decreasing in government assets, it varies too. These outcomes contrast sharply with those in a corresponding complete market benchmark like Lucas and Stokey's, where both debt and tax rates would be constant sequences, and with those in the incomplete markets economy of Aiyagari et al., where government assets approach levels that allow the limiting tax rate to be zero and the tail allocation to be first-best.

With more structure on the payoff vector, part 2 of theorem 3 shows that there is an inward drift to government assets: the sequence of Lagrange multipliers on the sequence of implementability constraints forms a sub (or super) martingale in regions with low (or high) debt. The envelope theorem links the dynamics of the multiplier to the dynamics of government debt. The concavity of the value function implies mean reversion for government debt. Mean reversion is particularly stark when $P(s) \in \mathcal{P}^*(s)$: here government debt converges to the constant B^* appearing in definition (20) of the set $\mathcal{P}^*(s)$.[14]

In part 3 of theorem 3 where $P(s) \in \mathcal{P}^*$, after government assets have converged to B^*, the government perfectly hedges fluctuations in its net-of-interest deficit. Whether the government eventually holds assets or owes debt is determined by the sign of the covariance of $P(s)$ with $g(s)$. Keeping the tax rate and therefore tax revenues constant, the government must finance a higher primary deficit when it gets a positive expenditure shock. If the asset returns more when government expenditures are high, the asset is a good hedge. In this situation, the government optimally holds positive assets and uses high returns on its holding to finance its net-of-interest deficit. On the other hand, if payoffs on the asset are lower when the government's net-of-interest deficit is high, then owing government debt is useful because it helps lower the government's interest burden in the face of adverse government expenditure shocks., after government assets have converged to B^*, the government perfectly hedges fluctuations in its net-of-interest deficit. Whether the government holds assets or owes debt is determined by the sign of the covariance of $P(s)$ with $g(s)$. Keeping the tax rate and therefore tax revenues constant, the government must finance a higher primary deficit when it gets a positive expenditure shock. If the asset returns more when government expenditures are high, the asset is a good hedge. In this situation, the government optimally holds positive assets and uses high returns on its holding to finance its net-of-interest deficit. On the other hand, if payoffs on the asset are lower when the government's net-of-interest deficit is high, then owing government debt is useful because it helps lower the government's interest burden in the face of adverse government expenditure shocks.

To acquire more information about the invariant distribution of government debt when the payoff vector $P(s)$ is close to the set $\mathcal{P}^*$, in theorem 4 we linearize the law of motion for government assets to approximate the mean and variance of the invariant distribution of government assets. We employ an orthogonal decomposition of an arbitrary $P(s) \notin \mathcal{P}^*$, namely,

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

where $P^*(s) \in \mathcal{P}^*$ and $\hat{P}(s)$ is orthogonal to $g(s)$. We carefully choose the point $P^*(s)$ about which we take a linear approximation, namely, a closest (in an l_2 sense) complete market economy that serves as the steady state of an economy for some $P(s) \in \mathcal{P}^*(s)$. Exploiting the structure of these approximate laws of motion allows us to obtain bounds on the standard deviation of

[14]Thus, the limiting allocation matches the allocation in a particular Lucas and Stokey economy with constant government debt and taxes; however, the level and the sign of long-run government debt is determined by the joint properties of shocks and payoffs rather than by the initial government debt, as it is in the Lucas and Stokey model.

government assets in the ergodic distribution and also the rate at which the mean asset level converges, a rate that can be expressed in terms of primitives in the form of the joint distribution of shocks and payoffs.

Theorem 4 *Under a first order approximation of dynamics around $P^*(s)$, the ergodic distribution of government assets has the following properties:*

- ***Mean:*** *The ergodic mean is B^* and thus equals the steady state level of government assets of an economy with payoff vector $P^*(s)$.*
- ***Variance:*** *The ergodic coefficient of variation of government assets B is*

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

- ***Convergence rate:*** *The speed of convergence to the ergodic distribution is*

$$\frac{\mathbb{E}_{t-1}(B_t - B^*)}{(B_{t-1} - B^*)} = \frac{1}{1 + |\text{cov}(P(s), g(s))|}.$$

We relegate the proof to appendix A.5, where we describe how we t take a first-order Taylor approximation to the decision rules and laws of motion for the state variables of our economy around complete market counterparts associated with $P^* \in \mathcal{P}^*$. An important feature is that the point around which the approximation is taken is not a deterministic steady state. The appendix also describes the accuracy of the approximation.

Theorem 3 established that when the payoff vector is $P^*(s)$ and the government's assets equal B^*, the planner wants the government to keep the tax rate constant and thereby perfectly hedge fluctuations in its net-of-interest deficit by using fluctuations in the earnings $P^*(s)B^*$ from its portfolio. When $P(s) \notin \mathcal{P}^*$, the incompleteness of markets prevents complete hedging, so shocks are only imperfectly hedged with a combination of changes in the tax rate and the level of government debt. Theorem 4 asserts how deviations from $\mathcal{P}^*$ map into larger variances for government debt and the tax rate under the ergodic distribution. Figure 1 illustrates how the ergodic distribution of government debt and the tax rate spread as we exogenously alter the covariance of $P(s)$ with $g(s)$. **Anmol and David XXXXX: was this graph generated from the approximation or from the more accurate approximation computed via the original hard line projection method?**

5.2 Heterogeneous agent economy with quasilinear preferences

We now alter the the section 5.1 economy by no longer restricting transfers to be zero and by adding a second agent who has zero productivity:

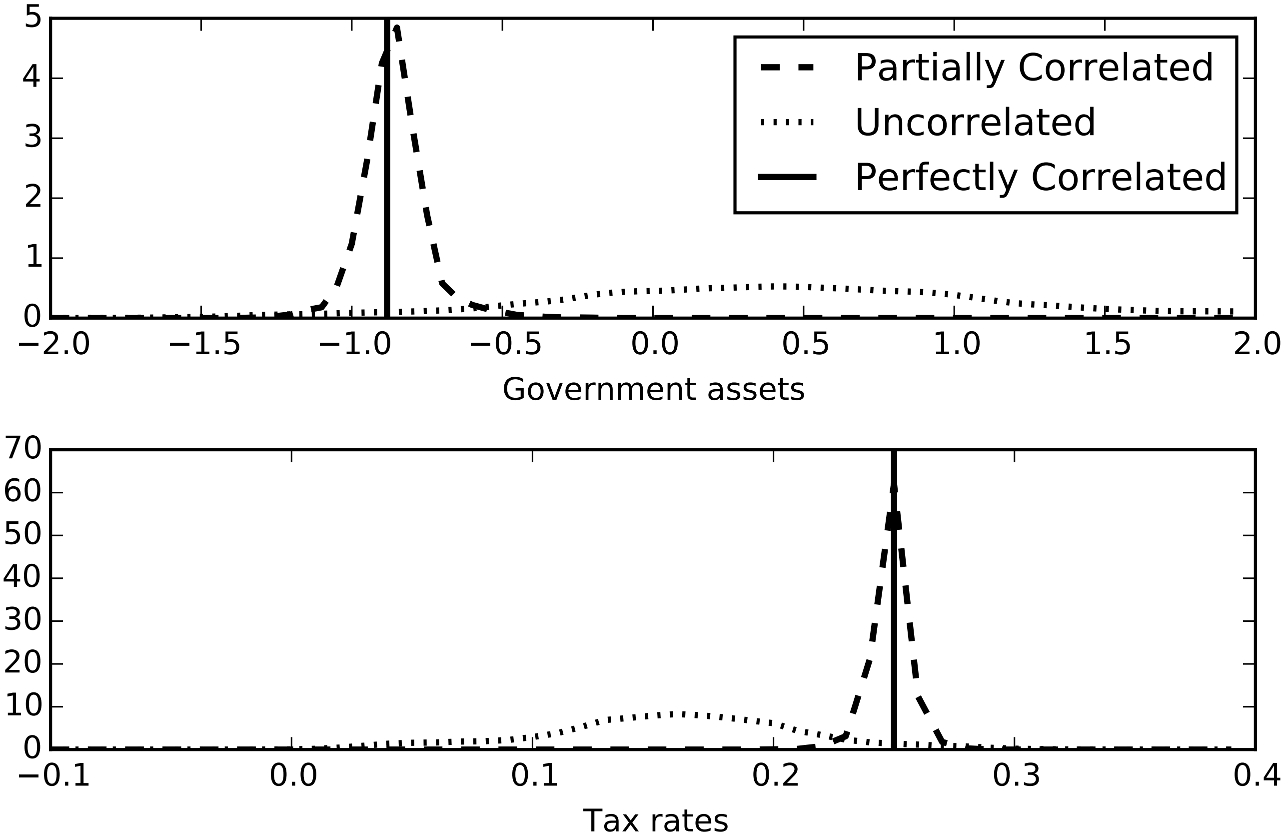


Figure 1: Ergodic distribution for government assets B_t and the labor tax rate τ_t in the representative agent quasilinear economy for three different asset payoff vectors$P(s)$.

Assumption 3 $\theta_1 > \theta_2 = 0$ and $c_{2,t} \geq 0$.

We use the freedom granted us by the Ricardian equivalence result summarized in theorem 1 and corollary 1 of section 3 to normalize the assets of the unproductive agent to zero throughout this section. In light of the zero productivity of a type 2 agent, this means that $c_{2t} = T_t$. Requiring that $c_{2t} \geq 0$ is an easy way to make transfers costly to the Ramsey planner for some values of the Pareto weights.

Theorem 5 *Let $(\omega, n) \in [0,1] \times [0,1]$ be the Pareto weight and mass assigned to the productive type 1 agent. Assume that $n < \frac{\gamma}{1+\gamma}$. The optimal tax rate, transfer, and government asset policies $\{\tau_t, T_t, B_t\}$ are characterized as follows:*

1. *For $\omega \geq n\left(\frac{1+\gamma}{\gamma}\right)$, $T_t = 0$. The optimal tax rate, government debt policy is same as in the representative agent economy studied in Theorems 3 and 4.*

2. *For $\omega < n\left(\frac{1+\gamma}{\gamma}\right)$, suppose that $\min_s\{P(s)\} > \beta$. There exist a $\mathcal{B}(\omega)$ satisfying $\mathcal{B}'(\omega) > 0$ and a $\tau^*(\omega)$ such that*

   (a) If $B_- > \mathcal{B}(\omega)$

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

   (b) If $B_- \leq \mathcal{B}(\omega)$, the Ramsey policies depend on the structure of $P(s)$.

   i. If $P(s) \notin \mathcal{P}^*$

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

   ii. If $P(s) \in \mathcal{P}^*$, we have two cases depending on B_-

   A. If $B_- \leq B^*$

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

   B. If $\mathcal{B}(\omega) > B_- > B^*$

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

Anmol XXXXX: We need to add a sentence saying "We relegate the proof to appendix XXXXX.

For a Pareto weight on the productive agent above a threshold $\bar{\omega} = n\left(\frac{1+\gamma}{\gamma}\right)$ featured in part 1 of theorem 5, the Ramsey planner chooses always to set transfers to zero. They are just too costly. Positive transfers entail subsidizing the unproductive type 2 agents whose consumption he values too little. The $T_t - 0$ outcome makes the Ramsey plan in the theorem 5 two-type of agents economy be identical to the Ramsey plan for the representative agent economy of theorem 3. The type type 1 agent in this economy in effect becomes the representative agent of the theorem 3 economy.

Things change in part 2 theorem 5 with a more progressive $\omega < \bar{\omega}$ Ramsey planner. Fluctuating transfers now become a useful tool to hedge aggregate shocks. If the government begins with enough assets (as in part 2.a of theorem 5), the planner chooses an interior allocation in which *all* fluctuations in net-of interest deficits are always financed by fluctuating transfers T_t. Here the tax rate is constant and government assets remain at their initial level where the marginal welfare costs of labor taxes are equated to the marginal welfare costs of transfers. Part 2.b describes a setting in which $P(s) \notin \mathcal{P}^*$ and government assets too low to validate the part 2.a outcome. Here the planner eventually accumulates government assets until they reach the threshold $\mathcal{B}(\omega)$ at which the welfare costs of transfers are low enough that the government can keep the tax rate constant.

The government's decision under the conditions of part 2 to accumulate enough assets so that eventually it can use earnings on its assets together with fluctuating positive transfers to hedge fiscal shocks is reminiscent of outcomes in the Aiyagari et al. (2002) economy. There, with a representative agent and non-negativity constraints on transfers, the planner accumulated enough assets so that it could finance shocks with zero distortionary labor taxes while costlessly using fluctuating transfers to dispose of excess earnings on its asset holdings. With multiple agents, fluctuating transfers can may bring welfare costs that depend on the Ramsey's planner attitude about redistribution. Theorem 5 describes how Pareto weights that make the planner care more about the unproductive agent lower the marginal welfare costs of collecting revenue from labor taxes paid by the productive agent. The planner thus increases the labor tax rate, lowering the threshold level of assets that are required to finance all shocks by transfers.

6 More general preferences

The quasilinearity of utilities maintained throughout section 5 set the equilibrium return on the asset to $\beta^{-1}P(s)$. Adding curvature to utility from consumption makes the return on the asset endogenous even for a risk-free bond having a payoff vector $P(s) = 1$. With such curvature, relative marginal utilities of consumption become state variables in a recursive representation of a Ramsey plan. We present a pair of Bellman equations that express a recursive representation of the Ramsey problem that in section 4 we formulated in a space of sequences. Here and also in section 7, we use this Bellman equation to analyze several economies with utility functions that are strictly concave in consumption.

6.1 Two Bellman equations

Let $\boldsymbol{x} = \left(U_c^1\tilde{b}_1, ..., U_c^{I-1}\tilde{b}_I\right)$, $\boldsymbol{\rho} = \left(U_c^1/U_c^I, ..., U_c^{I-1}/U_c^I\right)$, and denote an allocation $a = \{c_i, l_i\}_{i=1}^I$. Following Kydland and Prescott (1980) and Farhi (2010), we split the Ramsey problem into a time-0 problem that takes $(\{\tilde{b}_{i,-1}\}_{i=1}^{I-1}, s_0)$ as state variables and a time $t \geq 1$ continuation problem that takes $(\boldsymbol{x}, \boldsymbol{\rho}, s_{\_})$ as state variables. There are two value functions, one that pertains to $t \geq 1$, another to $t = 0$.[15] As usual, we work backwards and describe the time $t \geq 1$ Bellman equation first, and then the time $t = 0$ Bellman equation.

For $t \geq 1$, let $V(\boldsymbol{x}, \boldsymbol{\rho}, s_{\_})$ be the planner's continuation value given $\boldsymbol{x}_{t-1} = \boldsymbol{x}, \boldsymbol{\rho}_{t-1} = \boldsymbol{\rho}, s_{t-1} = s_{\_}$. It satisfies the Bellman equation

$$V(\boldsymbol{x}, \boldsymbol{\rho}, s_{\_}) = \max_{a(s), x'(s), \rho'(s)} \sum_s \pi(s|s_{\_}) \left(\left[\sum_i \omega_i U^i(s) \right] + \beta V(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s), s) \right) \tag{22}$$

where the maximization is subject to

$$U_c^i(s)\left[c_i(s) - c_I(s)\right] + x_i'(s) + \left(U_l^i(s) l_i(s) - U_c^i(s) \frac{U_l^I(s)}{U_c^I(s)} l_I(s) \right) = \frac{x P(s|s_{\_}) U_c^i(s)}{\beta \mathbb{E}_{s_{\_}} P \boldsymbol{U}_c^i} \text{ for all } s, i < I \tag{23a}$$

$$\frac{\mathbb{E}_{s_{\_}} P \boldsymbol{U}_c^i}{\mathbb{E}_{s_{\_}} P \boldsymbol{U}_c^I} = \rho_i \text{ for all } i < I \tag{23b}$$

$$\frac{U_l^i(s)}{\theta_i(s) U_c^i(s)} = \frac{U_l^I(s)}{\theta_I(s) U_c^I(s)} \text{ for all } s, i < I \tag{23c}$$

$$\sum_i n_i c_i(s) + g(s) = \sum_i n_i \theta_i(s) l_i(s) \quad \forall s \tag{23d}$$

$$\rho_i'(s) = \frac{U_c^i(s)}{U_c^1(s)} \text{ for all } s, i < I \tag{23e}$$

$$\sum_{i<I} \frac{x_i(s)}{U_c^i(s)} \text{ is bounded} \tag{23f}$$

Constraints (23b) and (23e) imply (8). The definition of x_t and constraints (23a) together imply equation (11) scaled by U_c^i.

Next we describe the Bellman equation pertinent for $t = 0$. Let $V_0\left(\{\tilde{b}_{i,-1}\}_{i=1}^{I-1}, s_0\right)$ be the value to the planner at $t = 0$, where $\tilde{b}_{i,-1}$ denotes initial debt and we retain the normalization $R_0 = \beta^{-1} P(s_0)$. It satisfies the Bellman equation

$$V_0\left(\{\tilde{b}_{i,-1}\}_{i=1}^{I-1}, s_0\right) = \max_{a_0, x_0, \rho_0} \sum_i \omega_i U^i(c_{i,0}, l_{i,0}) + \beta V\left(x_0, \rho_0, s_0\right) \tag{24}$$

[15] The time inconsistency of an optimal policy manifests itself in there being distinct value functions and Bellman equations at $t = 0$ and $t \geq 1$. For the quasilinear cases in sections 5, the Ramsey plans are time consistent.

where the maximization is subject to

$$U^i_{c,0}\left[c_{i,0}-c_{I,0}\right]+x_{i,0}+\left(U^i_{l,0}l_{i,0}-U^i_{c,0}\frac{U^1_{l,0}}{U^I_{c,0}}l_{I,0}\right)=\beta^{-1}P(s_0)U^i_{c,0}\tilde{b}_{i,-1}\text{ for all }i<I \tag{25a}$$

$$\frac{U^i_{l,0}}{\theta_{i,0}U^i_{c,0}}=\frac{U^I_{l,0}}{\theta_{I,0}U^{I,0}_{c}}\text{ for all }i<I \tag{25b}$$

$$\sum_i n_i c_{i,0}+g_0=\sum_i n_i\theta_{i,0}l_{i,0} \tag{25c}$$

$$\rho_{i,0}=\frac{U^i_{c,0}}{U^I_{c,0}}\ \forall\ i<I \tag{25d}$$

A tell-tale sign of the time inconsistency of the optimal plan is that (23b) is absent from the time 0 problem.

6.2 Eventual complete hedging with binary shocks

In this subsection, we present a class of economies with binary IID shocks in which the Ramsey planner can achieve complete hedging even when preferences are not quasilinear in consumption. Comovements of asset return with exogenous shocks govern the government's incentive to accumulate assets, an outcome reminiscent of the section 5.2 economy with quasilinear preferences. However, now the comovement of asset returns is endogenous.

For a given state $(\boldsymbol{x},\boldsymbol{\rho},s_{\_})$, let $\Psi\left(s;\boldsymbol{x},\boldsymbol{\rho},s_{\_}\right)=\left(x'\left(s\right),\rho'\left(s\right)\right)$ solve (22) so that $\Psi\left(s;\boldsymbol{x},\boldsymbol{\rho},s_{\_}\right)$ is the law of motion for the state variables under a Ramsey plan at $t\geq 1$.

Definition 4 *A steady state satisfies* $\left(\boldsymbol{x}^{SS},\boldsymbol{\rho}^{SS}\right)=\Psi\left(s;\boldsymbol{x}^{SS},\boldsymbol{\rho}^{SS},s_{\_}\right)$ *for all* $s,s_{\_}$.

In a steady state, the ratios of marginal utilities $\rho_i=U^i_c(s)/U^I_c(s)$ and marginal utility adjusted net assets x_i are constant; this means that the continuation allocation depends only on s_t and not on the history s^{t-1}.[16]

A competitive equilibrium allocation $\{c_i(s),l_i(s)\}_i$ associated with a choice for $\{\tau(s),\boldsymbol{\rho}(s)\}$ is determined by equations (23c), (23d) and (23e). We construct a set of economies having steady states. Denote $U(\tau,\boldsymbol{\rho},s)$ as the value of that competitive equilibrium allocation to a planner with Pareto weights $\{\omega_i\}_i$:

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

[16] History dependence is entirely intermediated through variation of $(\{x_i,\rho_i\}_i)$.

As before, define $Z_i(\tau, \rho, s)$ as

$$Z_i(\tau, \boldsymbol{\rho}, s) = U_c^i(s)c_i(s) + U_l^i(s)l_i(s) - \rho_i(s)\left[U_c^I(s)c_I(s) + U_l^I(s)l_I(s)\right].$$

When shocks are IID, the Ramsey policy solves the following Bellman equation in $\boldsymbol{x}(s^{t-1}) = \boldsymbol{x}, \boldsymbol{\rho}(s^{t-1}) = \boldsymbol{\rho}$

$$V(\boldsymbol{x}, \boldsymbol{\rho}) = \max_{\tau(s), \boldsymbol{\rho}'(s), \boldsymbol{x}'(s)} \sum_s \pi(s)\left[U(\tau(s), \boldsymbol{\rho}'(s), s) + \beta V(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s))\right] \tag{26}$$

where the maximization is subject to the constraints

$$Z_i(\tau(s), \boldsymbol{\rho}'(s), s) + x_i'(s) = \frac{x_i \beta^{-1} P(s) U_c^i(\tau(s), \boldsymbol{\rho}'(s), s)}{\mathbb{E} P U_c^i(\tau, \rho)} \text{ for all } s, i < I, \tag{27}$$

$$\sum_s \pi(s) P(s) U_c^1(\tau(s), \boldsymbol{\rho}'(s), s)(\rho_i'(s) - \rho_i) = 0 \text{ for } i < I. \tag{28}$$

Constraint (28), which rearranges constraint (23b), implies that $\rho(s)$ is a risk-adjusted martingale.

Our next job is to study conditions that render the first-order necessary conditions and feasibility as compressed into (27) and (28) consistent with the requirement that the implied law of motion for the state variables is a steady state associated with definition 4.[17]

Lemma 1 *When utility is strictly concave in consumption, $\|S\| = 2$ is necessary for a steady state to exist generically.*

Proof. Let $\pi(s)\mu_i(s)$ and λ_i be Lagrange multipliers on constraints (27) and (28). Imposing the restrictions $x_i'(s) = x_i$ and $\rho_i'(s) = \rho_i$, at a steady state $\{\mu_i, \lambda_i, x_i, \rho_i\}_{i=2}^N$ and $\{\tau(s)\}_s$ are determined by the following equations:

$$Z_i(\tau(s), \boldsymbol{\rho}, s) + x_i = \frac{\beta^{-1} P(s) x_i U_c^i(\tau(s), \boldsymbol{\rho}, s)}{\mathbb{E} U_c^i(\tau, \rho)} \text{ for all } s, i \geq 2, \tag{29a}$$

$$U_\tau(\tau(s), \boldsymbol{\rho}, s) - \sum_i \mu_i Z_{i,\tau}(\tau(s), \boldsymbol{\rho}, s) = 0 \text{ for all } s, \tag{29b}$$

$$U_{\rho_i}(\tau(s), \boldsymbol{\rho}, s) - \sum_j \mu_j Z_{j,\rho_i}(\tau(s), \boldsymbol{\rho}, s) + \lambda_i P(s) U_c^I(\tau(s), \boldsymbol{\rho}, s) - \lambda_i \beta \mathbb{E} P(s) U_c^I(\tau(s), \boldsymbol{\rho}(s), s) = 0. \text{ for all } s, i < I \tag{29c}$$

When the shock s takes only two values, (29) is a square system of $4(I-1)+2$ equations in $4(I-1)+2$ unknowns $\{\mu_i^{SS}, \lambda_i^{SS}, x_i^{SS}, \rho_i^{SS}\}_{i=1}^{I-1}$ and $\{\tau^{SS}(s)\}_s$. For $|S| \geq 3$, there are more equations than unknowns. ■

[17] Appendix A.8 discusses second-order conditions that ensure these policies are optimal.

At a steady state, outcomes resemble those in the complete market economy of Werning (2007). The tax rate and transfers both depend only on the current realization of shock s_t. Arguments of Werning can be adapted to show that the tax rate is constant when preferences have the CES form $c^{1-\sigma}/(1-\sigma) - l^{1+\gamma}/(1-\gamma)$, and also that fluctuations in the tax rate are very small when preferences take forms consistent with the existence of balanced growth.

6.3 A two-agent example

Since (29) is a non-linear system, in general existence of a solution can be verified only numerically. Here we provide a simple example with risk averse agents in which the existence of a root of (29) can be established analytically. This example will allow us to identify forces also present in the quasilinear economy of section 5.2 and the more general economies to be analyzed with numerical methods in section 7.

Consider an economy consisting of two types of households with $\theta_{1,t} > \theta_{2,t} = 0$ and common one-period utilities $\ln c - \frac{1}{2}l^2$. The shock s takes two values$\{s_L, s_H\}$ that are i.i.d across time. We assume that $g(s) = g$ for all s, and $\theta_1(s_H) > \theta_1(s_L)$.[18]

Theorem 6 *Suppose that $g < \theta(s)$ for all s. Let $R(s|s_-)$ be the return on the traded asset.*

1. ***Countercyclical returns.*** *If $P(s_H) = P(s_L)$, then there exists a steady state in which $B^{SS}(s) > 0$, $R^{SS}(s_H|s_-) < R^{SS}(s_L|s_-)$.*

2. ***Procyclical interest rate.*** *There exists a pair $\{P(s_H), P(s_L)\}$ implying a steady state with $B^{SS}(s) < 0$ and $R^{SS}(s_H|s_-) > R^{SS}(s_L|s_-)$.*

In both cases, the tax rate $\tau(s) = \tau^{SS}$ is independent of s.

Normalizing the assets of the unproductive agent to zero let us interpret B as the government's assets. Besides establishing existence of a steady state, theorem 6 isolates the procyclical or countercyclical comovement of the return on the asset as a predictor of the sign of government assets under a Ramsey plan.

Theorem 6 points to fluctuations in inequality and fluctuations in the return on the asset as two main forces governing determine the dynamics of the tax rate and the government's stock of the asset. If the asset return process were constant, n response to an adverse shock in the form of a fall in θ_1 the government could either increase the tax rate τ or it could decrease

[18] The restriction that expenditures are constant and productivities are stochastic can be relaxed and we obtain very similar results

transfers T. Both responses are distorting, but for different reasons. Increasing the tax rate increases distortions because the deadweight loss is convex in the tax rate, as in Barro (1979). The Ramsey planner copes with this distortion in the same way that it does in representative agent economies, namely, by trying to smooth the distorting tax rate. But in a heterogeneous agent economy like ours, adjusting transfers T is also costly. Starting from $B = 0$ (which means that agents' asset holdings are identical because we normalized by setting $b_2 = 0$), a decrease in transfers disproportionately adversely 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, giving rise to a cost not present in a representative agent economy.

The government can reduce the costs of attaining its objectives about redistribution by choosing tax rate policies that make the net assets of the high-skilled agent decrease over time. That makes the two agents' after-tax and after-interest income become closer, allowing decreases in transfers to have smaller effects on inequality in marginal utilities. 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.[19] This pushes B to be positive in the long run.

Turning now to the second force, the return on the asset generally fluctuates with shocks. Parts 1 and 2 of theorem 6 isolate forces that drive those fluctuations. Consider again the example of a decrease in the productivity of high-skilled agents. If the tax rate τ is left unchanged, since g is constant, the government requires extra revenues. But suppose that the return on the asset increases whenever θ_1 decreases, as happens, for example in part 1 of theorem 6 with a risk free bond. If the government holds positive assets, its earnings from those assets increase. So holding assets allows higher income from assets to offset some of the government's revenue losses from taxes on labor. The situation reverses if the return on the asset falls at times of increased need for government revenues coming from higher net-of interest deficits, as in part 2 of theorem 6, so the steady state sees the government's owing debt.

In the long run, the government's debt depends on the balance of the two forces: (1) inequality distortions push the government asset position B to be positive, and (2) hedging motives that can also push B positive as in part 1 theorem 6, but that for sufficiently procylical returns can push government assets B to be negative in the long run as in part 2 of the theorem.

These outcomes have counterparts in the representative agent quasilinear economy studied in section 5. There, quasilinearity of preferences allowed us to characterize how the covariance

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

of the asset returns and exogenous shocks affected the sign (and level) of government assets through expression (21). In parts 1 and 2 of theorem 6, with binary shocks, altering the gap $P(s_H) - P(s_L)$ implies a such variation in asset returns. Despite the endogeneity of the key covariance in the model of this section, the underlying forces remain those prevailing more immediately in section 5.

6.4 Stability

To say some things about the long-run behavior of government debt in more general economies, we now apply an extension of the Theorem 4 approximation method. In the quasilinear case, we obtained an analytical characterization. Here, for more general specifications, we use a numerical convergence criterion to establish local stability of a steady state over a range of parameter values. **Anmol XXXXX: please describe the dimensions of more generality – e.g., persistence, curvature, number of agents – in the preceding sentence.**

As before, let $\pi(s)\mu_i(s)$ and λ_i be Lagrange multipliers on constraints (27) and (28), respectively. In Appendix A.8 we show that the history-dependent Ramsey policies (they are sequences of functions of s^t) can be represented recursively in terms of $\{\boldsymbol{\mu}(s^{t-1}), \boldsymbol{\rho}(s^{t-1})\}$ and s_t. A recursive representation of an optimal policy can be linearized around steady state values of the state variables $(\boldsymbol{\mu}, \boldsymbol{\rho})$.[20] Let $\hat{\Psi}_t = \begin{bmatrix} \boldsymbol{\mu}_t - \boldsymbol{\mu}^{SS} \\ \boldsymbol{\rho}_t - \boldsymbol{\rho}^{SS} \end{bmatrix}$ be deviations from a steady state. Construct a linear approximation

$$\hat{\Psi}_{t+1} = C(s_{t+1})\hat{\Psi}_t \tag{30}$$

whose coefficients $C(s)$ are functions of the shock s. Let $\mathbb{E}C(s)$ be the mathematical expectation of $C(s)$. The next theorem describes features of $\mathbb{E}C(s)$ that control whether this linear system converges to zero in probability.

Theorem 7 *If the (real parts) of the eigenvalues of $\mathbb{E}C(s)$ are less than 1, system (30) converges to zero in mean. For large t, the conditional variance of $\hat{\Psi}$, denoted by $\Sigma_{\Psi,t}$, is governed by*

$$vec(\Sigma_{\Psi,t}) = \widehat{C} vec(\Sigma_{\Psi,t-1}),$$

where $\widehat{C}$ is a $(2I-2) \times (2I-2)$ matrix. In addition, if the (real parts) of the eigenvalues of $\widehat{C}$ are less than 1, system (30) *converges in probability.*

[20] One could in principle look for a solution in state variables $\left(\boldsymbol{x}(s^{t-1}), \boldsymbol{\rho}(s^{t-1})\right)$. For $I = 2$ with $\{\theta_i(s)\}$ different across agents, this would give identical policies and a map that 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 perturbations in $(\boldsymbol{\mu}, \boldsymbol{\rho})$.

The dominant eigenvalue of $\mathbb{E}C(s)$ is informative not only about whether the system is locally stable but also about the rate at which the steady state is approached. The half-life of convergence to the steady state is $\frac{\log(0.5)}{\|\iota\|}$, where $\|\iota\|$ is the absolute value of the dominant eigenvalue.

In work described in Appendix A.9, we have applied Theorem 7 to verify local stability for many examples. The typical finding there is that the steady state is stable but that convergence is slow. The rates of convergence are increasing in the strength of covariance of the return on the asset and aggregate shocks that affect the net-of-interest government deficit. We return to this feature when we study low frequency components of the Ramsey allocation for the quantitative example to be studied in section 7.

7 Numerical example

In sections 5 and 6, we studied steady states as a way of summarizing the asymptotic behavior of Ramsey allocations and the forces that shape the asymptotic distribution of government and private assets. In this section, we use a calibrated version of the economy a) to revisit the magnitude of these forces; and b) to study optimal policy responses at business cycle frequencies when the economy is possibly far away from a (stochastic) steady state. We choose shocks and initial conditions to match stylized facts from the recent recession in US. The numerical calculations use methods adapted from Evans (2014) and described in the Appendix A.10.

7.1 Calibration

We assume five types of agents of equal measures with preferences $u(c,l) = \frac{c^{1-\sigma}}{1-\sigma} - \frac{l^{1+\gamma}}{1+\gamma}$. These agents stand in for the $90^{\text{th}}, 75^{\text{th}}, 50^{\text{th}}, 25^{\text{th}}$, and 10^{th} quantiles of the US wage distribution.[21]

Let $\mathcal{Q}(i)$ be the quantile of agent i. We assume i.i.d aggregate shocks ϵ_t that affect both the labor productivities of all agents $\{\theta_{i,t}\}_{i=1}^{I}$ and the payoff p_t of the single asset:

$$\log\theta_{i,t} = \log\bar{\theta}_i + \epsilon_t[1 + (.9 - \mathcal{Q}(i))m] \tag{31a}$$

$$p_t = 1 + \chi\epsilon_t \tag{31b}$$

Following Autor et al. (2008),we set average productivities $\{\bar{\theta}_i\}_{i=1}^{N}$ to match quantiles of average weekly earnings of full time wage and salary earners from the Current Population Survey (CPS).

[21] Setting $I = 5$ allows us ample heterogeneity in wealth and earnings. We have verified that the main qualitative and quantitative insights are unchanged when we have more than five types. Our equilibrium approximation method is practical with a much larger number of types than 5.

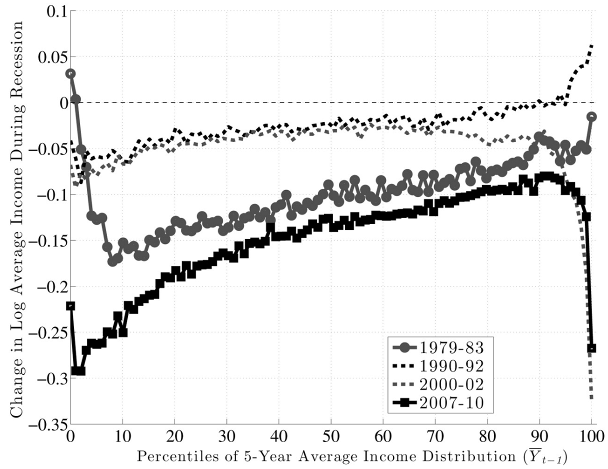


Figure 2: Change in log average earnings during recessions, prime-age males from Guvenen et al. (2014)

The parameter m allows us to generate recessions associated with different falls in income for different types of agents. We calibrate m to match facts reported by Guvenen et al. (2014). Figure 7.1 (adapted from Guvenen et al.) reports that in the latest US recession the fall in income for agents in the first decile of earnings was about three times that experienced by the 90th percentile. Furthermore, between the 10th and the 90th percentiles, the change in the percentage drop in earnings was almost linear. From these facts we infer a slope $m = \frac{1.5}{0.8}$.

The parameter χ captures the ex-post comovement in returns on government assets and aggregate shocks. Our model is silent about the source of these comovements. In the data, they could compe from variations in real payoffs due inflation, interest rate risk for longer maturity bonds, or defaults. For the purpose of our numerical exercise, we use US data on inflation and interest rates of longer maturities bonds to calibrate χ. We calibrate the comovement in the following way. Let $q_t^{(n)}$ be the log price of a nominal bond of maturity n. We can define real holding period returns $r_{t,t+1}^{(n)}$ as

$$r_{t,t+1}^{(n)} = q_{t+1}^{(n-1)} - q_t^{(n)} - \pi_{t+1}$$

With the transfromation $y_t^{(n)} : -\frac{1}{n} q_t^{(n)}$ we can express $r_{t,t+1}^{(n)}$ as follows:

$$r_{t,t+1}^{(n)} = \underbrace{y_t^{(n)}}_{\text{Ex-ante part}} - (n-1)\left[\underbrace{\left(y_{t+1}^{(n)} - y_t^{(n)}\right)}_{\text{Interest rate risk given } n} + \underbrace{\left(y_{t+1}^{(n-1)} - y_{t+1}^{(n)}\right)}_{\text{Term structure risk}}\right] - \underbrace{\pi_{t+1}}_{\text{Inflation risk}}$$

In our model, the holding period returns are given by $\log\left[\frac{p_{t+1}}{q_t}\right]$ and $q_t = \frac{\beta \mathbb{E}_t u_{c,t+1} P_{t+1}}{u_{c,t}}$. Note that p_{t+1} allows us to captures ex-post fluctuations in returns to the government's debt portfolio coming from maturity and inflation.

Table 1 summarizes the comovement between labor productivity $\{\epsilon_t\}$ and bond prices $\{q_t^n\}$ for different maturities inferred from quarterly US data for the period 1952 to 2003. The table's first line reports the correlation between the ex post returns and labor productivities. In our baseline calibration, ϵ_t is i.i.d over time. Hence the parameter $\chi = \frac{\sigma_r}{\sigma_\epsilon} Corr(r, \epsilon)$. By averaging over different maturities we infer a value of $\chi = -0.06$. [22] Thus, payoffs are weakly countercyclical for US. Besides the results for the benchmark value of $\chi = -0.06$, the long simulations in section 7.2 include outcomes for a range of χ's from -1.0 to 1.0.

Maturity (n)	2yr	3yr	4yr	5yr
$Corr(\epsilon_{t+1}, r_{t,t+1}^{(n)})$	-0.11	-0.093	-0.083	-0.072
$Corr(\epsilon_{t+1}, r_{t,t+1}^{(n)} - ny_t^{(n)})$	0.00	-0.0463	-0.080	-0.091
$Corr(\epsilon_{t+1}, y_t^{(n)} - \pi_{t+1})$	-0.097	-0.086	-0.080	-0.073
$\frac{\sigma(r_{t+1}^n)}{\sigma(\epsilon_{t+1})}$	0.820	0.835	0.843	0.845

Table 1: Correlation between holding period returns and productivity

As for parameters of household preferences, we set $\sigma = 1$, $\gamma = 2$, which imply Frisch elasticity of labor supply of 0.5. We set the time discount factor $\beta = 0.98$, which implies the annual interest rate in an economy without shocks would be 2% per year.

We assume that the initial wealth is perfectly correlated with wages and calibrate the wealth distribution to get the relative quantiles as in Kuhn (2014) and Quadrini and Rios-Rull (2014). These papers document the quantiles of net worth for US households computed up to 2010 Survey of Consumer Finances.

For the Pareto weights and government expenditures, we use an optimal allocation in an economy without shocks to target a (pre-transfers, federal) expenditure output ratio of 12%, a tax rate of 23%, a ratio of transfers to gdp of 10%, and a government debt to gdp of 100%.

7.2 Long run outcomes

Figure 7.2 simulations of 2000 periods for the government debt to output ratio, the labor tax rate, and the transfers to output ratio for three values of $\chi \in \{-1.0, -0.06, 1.0\}$ in red, black, and blue, respectively. The three simulations start from the same initial conditions and all share the same sequence of underlying shocks.

[22] The second line of table 1 computes the correlation of labor productivity with the ex-post component of returns. For the shortest maturity, 3 month real tbill returns $Corr(\epsilon_{t+1}, y_t^{1qtr} - \pi_{t+1}) = -0.11$. These results together give us a range for χ of zero to negative -0.09. The numerical results are not sensitive to values of χ is this range.

Parameter	Value	Description
$\{\bar{\theta}_i\}$	{4.9,3.24,2.1,1.4,1}	Wages dispersion for {90,75,50,25,10} percentiles
γ	2	Average Frisch elasticity of labor supply of 0.5
β	0.98	Average (annual) risk free interest rate of 2%
m	$\frac{1.5}{.8}$	Heterogeneity in wage growth over business cycles
χ	-0.06	Covariance between holding period returns and labor productivity%
σ_e	0.03	vol of labor productivity
g	.13 %	Average pre-transfer expenditure-output ratio of 12 %

Table 2: Benchmark calibration

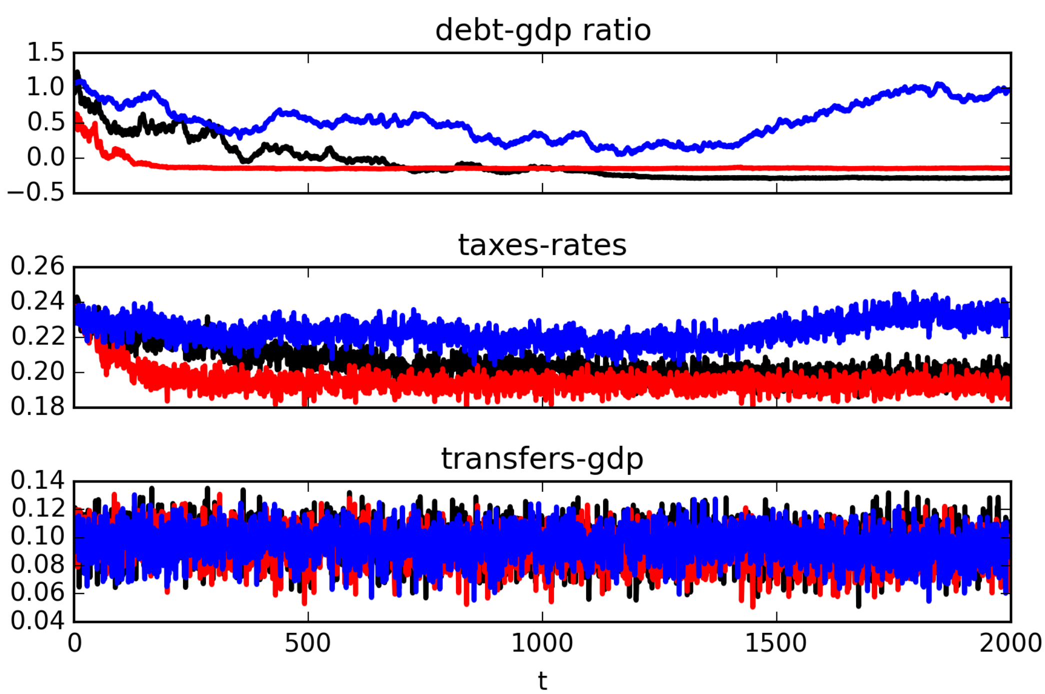


Figure 3: The red, black and blue lines plot simulations for a common sequence of shocks for values of $\chi = -1.0, -0.06, 1.0$ respectively

Two features emerge. Different values χ give rise to different locations of the long-run marginal distribution of government assets and also to different rates of convergence to that long-run distribution. A sufficiently positive χ generates lower payoffs in recessions relative to booms. In line with assertions of theorems 3 and 6, we see from the blue line that the government does not repay its initial debt during these 2000 periods. On the other hand, under the benchmark χ (black line) or when χ is negative (red line), the government accumulates assets.

In order to get a clearer picture of the speed of convergence, we plot paths of the conditional means for debt and the tax rate in figure 7.2. To explain how we generated these plots, let $B(s_{t+1}, \boldsymbol{x}_t, \boldsymbol{\rho}_t)$ be the Ramsey decision rules that generate the assets B of the government and let $\Psi\left(s_{t+1}; \boldsymbol{x}_t, \boldsymbol{\rho}_t\right)$ be the law of motion for the state variables for the Ramsey plan. For a given history, the conditional mean of government assets is:

$$B_{t+1}^{cm} = \mathbb{E}B(s_{t+1}, \boldsymbol{x}_t^{cm}, \boldsymbol{\rho}_t^{cm}) \tag{32a}$$

$$\boldsymbol{x}_t^{cm}, \boldsymbol{\rho}_t^{cm} = \mathbb{E}\Psi(s_t, \boldsymbol{x}_{t-1}^{cm}, \boldsymbol{\rho}_{t-1}^{cm}) \tag{32b}$$

Note how these conditional mean paths smooth the high frequency movements in the dynamics of the state variables but retain the low frequency drifts. As before, different lines correspond to different values of χ between -1.0 and 1.0 with the blue (red) lines representing positive (negative) values of χ. Thicker lines depict outcomes associated with larger values of χ. The figure shows that the speed of convergence is increasing and the magnitude of the limiting assets in decreasing in the strength of correlation between productivities and payoffs. This pattern confirms the approximation results characterized in theorem 4.

To verify the wide support of the ergodic distributions, we take the initial conditions at the end of the long simulation and subject the economy to a sequence of 100 periods of ϵ_t shocks that are 2 standard deviations below the mean. In figure 7.2 we see that given a sufficiently long sequence of negative productivity shocks the economy will eventually deviate significantly from its ergodic mean.

A further inference from the analysis of earlier sections was that government assets B in the steady state are decreasing in the redistributive motive of the government. We check this numerically here by changing Pareto weights. We parametrize the redistributive motive using α. The planner places evenly spaced Pareto weights from $0.2 + \alpha$ on the lowest productivity agent to $0.2 - \alpha$ on the highest productivity agent. Increasing α lowers concerns for redistribution. In figure 7.2 we plot mean of the government assets in the ergodic distribution as a function of α.

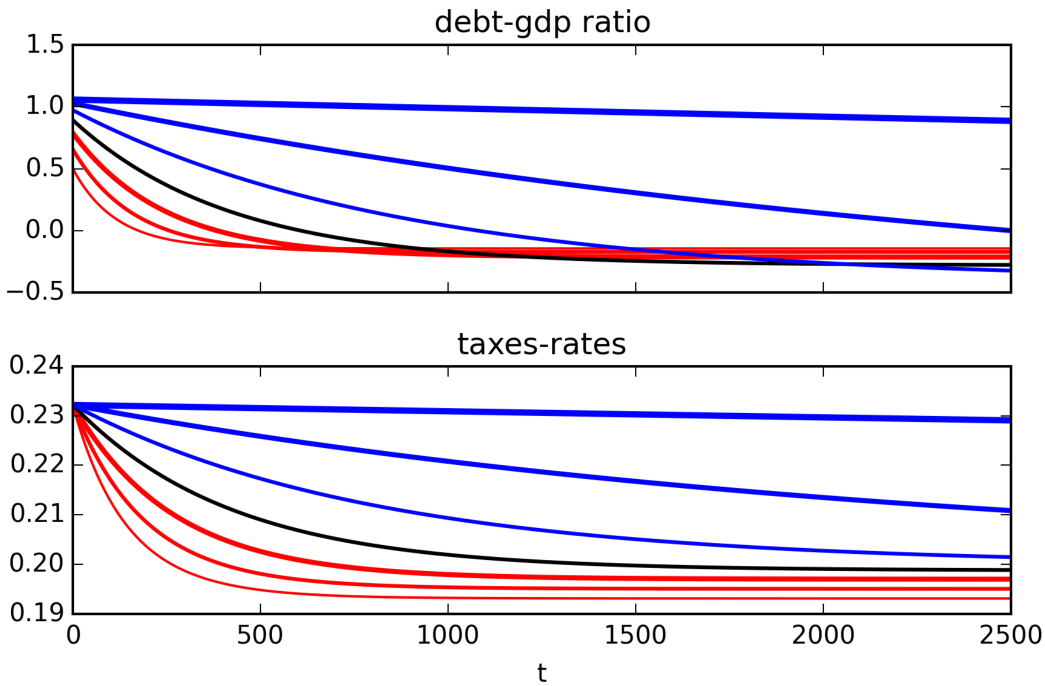


Figure 4: The plot shows conditional mean paths for different values of χ. The red (blue) lines have $\chi < 0$ ($\chi > 0$). The thicker lines represent larger values.

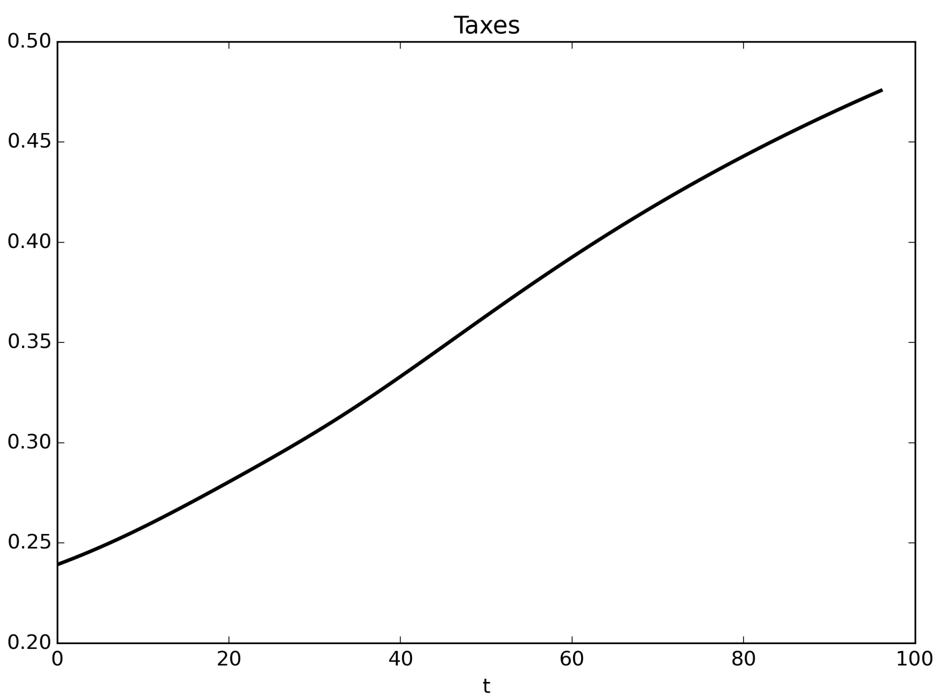


Figure 5: Taxes for a sequence of -1 s.d shocks to aggregate productivity of length 100

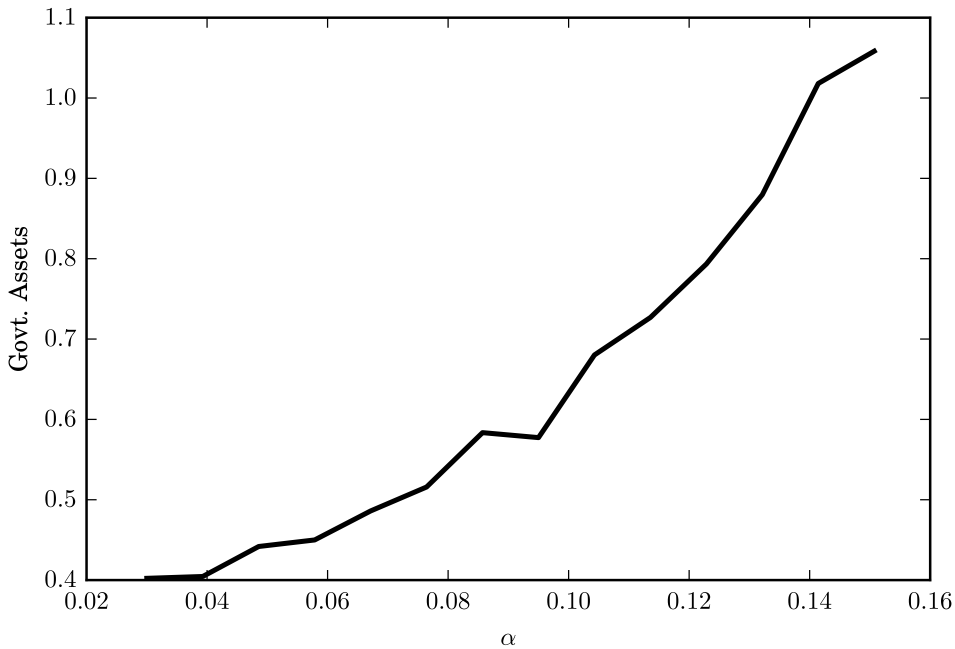


Figure 6: This plot shows long run assets of the government as a function of α which parametrizes the redistributive concern. Higher α represent planner's with relatively higher weights on productive agents.

7.3 Short run

The analysis of the previous subsection studied very low frequency components of a Ramsey plan. Here we focus on business cycle frequencies. In our setting, these higher frequency responses can conveniently be classified in terms of magnitudes of changes as we switch from "boom" to "recession," and the dynamics during periods when a recession or boom state persists. A recession is a negative -1.0 standard deviation realization for the ϵ_t process. Given the initial conditions and the benchmark calibration, the plots below trace the paths for debt, the tax rate, and transfers for a sequence of shocks that feature a recession of four periods from $t = 3$. Before and after this recession, the economy receives $\epsilon_t = 0$.

The main exercise here is to compute how the Ramsey tax rate, transfers, and government debt in recessions accompanied by larger inequality differ from those in a recession that affects all agents alike. Under the benchmark calibration, log wages for agent i are given by $\log \theta_i = \log \overline{\theta}_i + \epsilon[1 + (.9 - \mathcal{Q}(i))m]$. We decompose the total responses into a TFP only component by setting $m = 0$ and an inequality only component as follows:

$$\log \theta_i^{tfp} = \log \overline{\theta}_i + \epsilon$$

$$\log \theta_i^{ineq} = \log \overline{\theta}_i + \epsilon[(.9 - \mathcal{Q}(i))m]$$

Figure 7.3 plots impulse responses. The shaded region is the induced recession and the bold line captures the benchmark (total) response. The dashed (dotted) line reflects the TFP only (inequality) effect. In the benchmark, the government responds to an adverse shock by a making big increases in transfers, the tax rate, and government debt. However, without inequality shocks (dotted line), the government responds by decreasing transfers and increasing both debt and the tax rate, but by amounts an order of magnitude smaller than in the benchmark.

Next we average over sample paths of length 100 periods and report the volatility, autocorrelation, and correlation with exogenous shocks for the tax rate and transfers in table 3. We see that taxes are twice as volatile and that the correlation between transfers and productivities switches sign. This indicates how ignoring redistributive goals affect prescriptions for government policy in recessions.

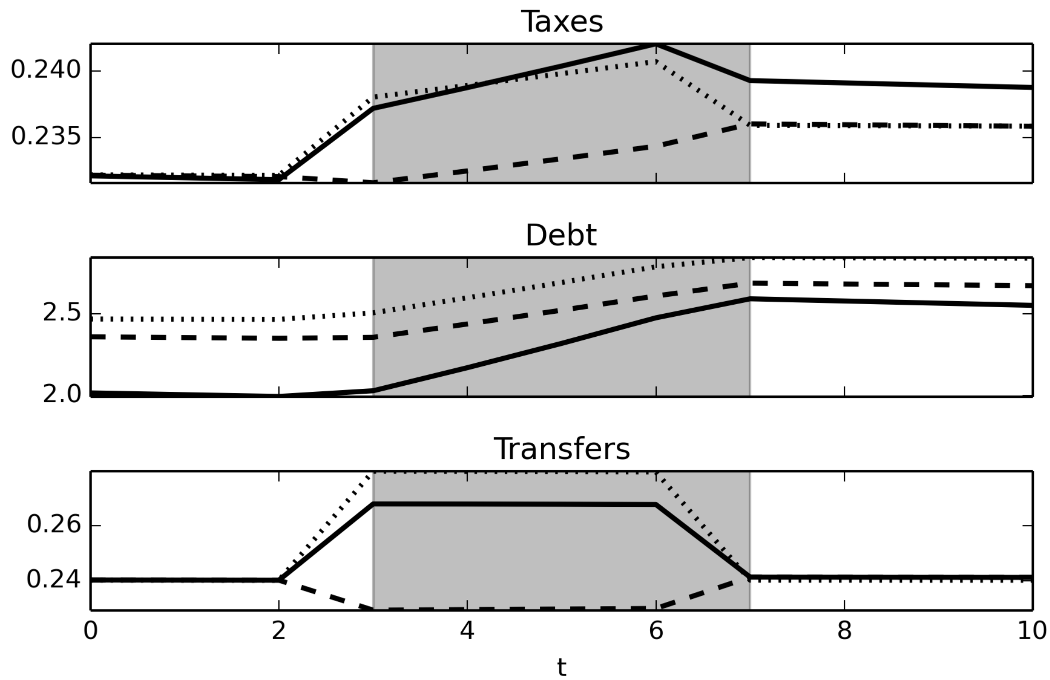


Figure 7: The bold line is the total response. The dashed (dotted) line reflects the only TFP (inequality) effect. The shaded region is the recession

Moments	Tfp	Tfp+Ineq
vol. of taxes	0.003	0.006
vol. of transfers	0.01	0.02
autocorr. in taxes	0.93	0.66
autocorr. in transfers	0.17	0.18
corr. of taxes with tfp	0.15	-0.63
corr. of transfers with tfp	0.99	-0.98

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

8 Conclusion

A Appendix

A.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.[23] In economies with exogenous borrowing constraints, agents' maximization problems include the additional constraints

$$b_{i,t} \geq \underline{b}_i \tag{33}$$

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

Definition 5 *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 (33),* $\{b_{i,t}\}_{i,t}$ *are bounded, and constraints (**??**), (5) and (3) are satisfied.*

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

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 theorem asserts that ad-hoc borrowing limits do not limit a government's ability to respond to aggregate shocks.[24]

Theorem 8 *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 $\hat{b}_{i,-1} = b_{-1}$.

[23] 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.

[24] See Yared (2012, 2013) who shows a closely related result.

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 (33). 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 provide some intuition for Theorem 8, 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 relative to an economy without borrowing constraints because a government might want to push some agents against their borrowing limits. When 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. We next construct an example without any shocks in which the government can achieve higher welfare by using borrowing constraints to improve its ability to redistribute. 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 a deterministic economy with $g_t = 0$, $\beta_t = \beta$ and $I = 2$. Further the utility function over consumption and labor supply $U(c, l)$ is separable in the arguments and satisfies the Inada conditions. The planners problem can then be written as the following sequence problem

$$\max_{\{c_{i,t}, l_{i,t}, b_{i,t}, R_t\}_t} \sum_{t=0}^{\infty} \beta^t \left[\alpha_1 U(c_{1,t}, l_{1,t}) + \alpha_2 U(c_{2,t}, l_{2,t})\right] \tag{34}$$

subject to

$$c_{2,t} + \frac{U_{l2,t} l_{2,t}}{U_{c2,t}} - \left(c_{1,t} + \frac{U_{l1,t} l_{1,t}}{U_{c1,t}} \right) + \frac{1}{R_t} (b_{2,t} - b_{1,t}) = b_{2,t-1} - b_{1,t-1} \tag{35a}$$

$$\frac{U_{l1,t}}{\theta_1 U_{c1,t}} = \frac{U_{l2,t}}{\theta_2 U_{c2,t}} \tag{35b}$$

$$c_{1,t} + c_{2,t} \leq \theta_1 l_{1,t} + \theta_2 l_{2,t} \tag{35c}$$

$$\left(\frac{U_{ci,t}}{U_{ci,t+1}} - \beta R_t \right) (b_{i,t} - \underline{b}_i) = 0 \tag{35d}$$

$$\frac{U_{ci,t}}{U_{ci,t+1}} \geq \beta R_t \tag{35e}$$

$$b_{i,t} \geq \underline{b}_i \tag{35f}$$

Where $\underline{b}_i$ is the exogenous borrowing constraint for agent i. We obtain equation (35a) by eliminating transfers from the budget equations of the households and using the optimality for labor supply decision. Equations (35d) and (35e) capture the inter-temporal optimality conditions modified for possbily binding constraints.

Let c_i^{fb} and l_i^{fb} be the allocation that solves the first best problem, that is maximizing equation (34) subject to (35c), and define

$$Z^{fb} = c_2^{fb} + \frac{U_{l2}^{fb} l_2^{fb}}{U_{c2}^{fb}} - \left(c_1^{fb} + \frac{U_{l1}^{fb} l_1^{fb}}{U_{c1}^{fb}} \right) \tag{36}$$

and

$$\tilde{b}_2^{fb} = \frac{Z^{fb}}{\frac{1}{\beta} - 1} \tag{37}$$

We will assume that the exogenous borrowing constraints satisfy $\underline{b}_2 = \underline{b}_1 + \tilde{b}_2^{fb}$. We then have the following lemma

Lemma 2 *If $\tilde{b}_2^{fb} > (<) 0$ and $b_{2,-1} - b_{1,-1} > (<) \tilde{b}_2^{fb}$ then the planner can implement the first best.*

Proof. We will consider the candidate allocation where $c_{i,t} = c_i^{fb}$, $l_{i,t} = l_i^{fb}$, $b_{i,t} = \underline{b}_i$ and interest rates are given by $R_t = \frac{1}{\beta}$ for $t \geq 1$. It should be clear then that equations (35b) and (35c) are satisfied as a property of the first best allocation. Equation (35d) is trivially satisfied since the agents are at their borrowing constraints. For $t \geq 1$ equations (35a) and (35e) are both satisfied by the choice of $R_t = \frac{1}{\beta}$ and the first best allocations. It remains to check that equation (35a) is satisfied at time $t = 0$ for an interest rate $R_0 < \frac{1}{\beta}$. At time zero the constraint is give by

$$Z^{fb} + \frac{1}{R_0} \tilde{b}_2^{fb} = b_{2,-1} - b_{1,-1} \tag{38}$$

The assumption that $b_{2,-1} - b_{1,-1} > (<)\tilde{b}_2^{fb}$ if $\tilde{b}_2^{fb} > (<)0$ then implies that

$$R_0 = \frac{\tilde{b}_2^{fb}}{b_{2,-1} - b_{1,-1} - Z^{fb}} < \frac{1}{\beta}$$

as desired. ■

This will improve upon the planners problem without exogenous borrowing constraints, as first best can only be achieved in this scenario when $b_{2,-1} - b_{1,-1} = \tilde{b}_2^{fb}$.

A.2 Proof of Theorem 1

Proof. Let

$$\hat{T}_t = T_t + \left(\hat{b}_{1,t} - b_{1,t}\right) - R_{t-1}\left(\hat{b}_{1,t-1} - b_{1,t-1}\right) \text{ for all } t \geq 0. \tag{39}$$

Given a tax policy $\left\{\tau_t, \hat{T}_t\right\}_t$, the allocation $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_{i,t}$ 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 $\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_{i,t}$ 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. ∎

A.3 Proof of Theorem 2

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 $\mathbf{x}(s^t)$ be a vector of n goods and let $\mathbf{p}(s^t)$ 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. Let $b(s^t)$ be the agent's asset holdings, and let $\mathbf{e}\left(s^t\right)$ be a stochastic vector of endowments.

Consumer maximization problem

$$\max_{\mathbf{x}_t, b_t} \sum_{t=0}^{\infty} \beta^t \Pr\left(s^t\right) U(\mathbf{x}\left(s^t\right)) \tag{40}$$

subject to

$$\mathbf{p}\left(s^t\right) \mathbf{x}\left(s^t\right) + q(s^t) b\left(s^t\right) = \mathbf{p}\left(s^t\right) \mathbf{e}\left(s^t\right) + P(s_t) b\left(s^{t-1}\right) \tag{41}$$

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

$$\begin{aligned} \mathbf{U}_x(s^t) &= U_1(s^t)\mathbf{p}(s^t) \\ \Pr\left(s^t\right) U_1\left(s^t\right) q(s^t) &= \beta \sum_{s^{t+1} > s^t} \Pr\left(s^{t+1}\right) U_1\left(s^{t+1}\right). \end{aligned} \tag{42}$$

Theorem 9 *Consider an allocation $\{\mathbf{x}_t, b_t\}$ that satisfies (41), (42) and $\{b_t\}_t$ is bounded. Then $\{\mathbf{x}_t, b_t\}$ is a solution to (40).*

Proof. The proof follows closely Constantinides and Duffie (1996). Suppose there is another budget feasible allocation $\mathbf{x} + \mathbf{h}$ that maximizes (40). Since U is strictly concave,

$$\begin{aligned} & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(\mathbf{x}_t + \mathbf{h}_t) - \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(\mathbf{x}_t) \\ \leq \; & \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \mathbf{U}_x(\mathbf{x}_t) \mathbf{h}_t \end{aligned} \tag{43}$$

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

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

Multiply by $\beta^t \Pr\left(s^t\right) U_1(s^t)$ to get:

$$\begin{aligned} \beta^t \Pr\left(s^t\right) U_1(s^t)\mathbf{p}(s^t)\mathbf{h}\left(s^t\right) &= \beta^t \Pr\left(s^t\right) U_1(s^t)\varphi(s^{t-1}) - q(s^t)\beta^t \Pr\left(s^t\right) U_1(s^t)\varphi(s^t) \\ &= \beta^t \Pr\left(s^t\right) U_1(s^t)\varphi(s^{t-1}) - \beta^{t+1} \sum_{s^{t+1} > s^t} \Pr\left(s^{t+1}\right) U_1\left(s^{t+1}\right) \varphi(s^t) \end{aligned}$$

where we used the second part of (42) in the second equality. Sum over the first T periods and use the first part of (42) to eliminate $\mathbf{U}_x(\mathbf{x}_t) = U_1(s^t)\mathbf{p}(s^t)$

$$\sum_{t=0}^{T} \beta^t \Pr\left(s^t\right) \mathbf{U}_x(\mathbf{x}_t)\mathbf{h}\left(s^t\right) = -\sum_{s^{T+1}>s^T} \beta^{T+1} \Pr\left(s^{T+1}\right) U_1\left(s^{T+1}\right) \varphi(s^T).$$

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

$$\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \Pr\left(s^t\right) \mathbf{U}_x(\mathbf{x}_t)\mathbf{h}\left(s^t\right) = 0.$$

Substitute this into (43) to show that $\mathbf{h}$ does not improve utility of consumer. ■

A.4 Proof of Theorem 3

Proof. The optimal Ramsey plan solves the following Bellman equation. Let $V(b_{-})$ be the maximum ex-ante value the government can achieve with debt b_{-}.

$$V(b_{-}) = \max_{c(s),l(s),b(s)} \sum_{s} \pi(s) \left\{ c(s) - \frac{l(s)^{1+\gamma}}{1+\gamma} + \beta V(b(s)) \right\} \tag{44}$$

subject to

$$c(s) + b(s) = l(s)^{1+\gamma} + \beta^{-1} P(s) b_{-} \tag{45a}$$

$$c(s) + g(s) \leq \theta l(s) \tag{45b}$$

Let $\bar{b} = -\underline{B}$

$$\underline{b} \leq b(s) \leq \bar{b} \tag{45c}$$

Let $\mu(s), \phi(s), \underline{\kappa}(s), \overline{\kappa}(s)$ be the Lagrange multipliers on the respective constraints. Part 1 of Theorem **3**

Lemma 3 *There exists a $\overline{b}$ such that $b_t \leq \overline{b}$. This is the natural debt limit for the government.*

Proof. As we drive μ to $-\infty$, the tax rate approaches a maximum limit, $\bar{\tau} = \frac{\gamma}{1+\gamma}$. In state s, the government surplus,

$$S(s, \tau) = \theta^{\frac{\gamma}{1+\gamma}} (1-\tau)^{\frac{1}{\gamma}} \tau - g(s),$$

which is maximized at $\tau = \frac{\gamma}{1+\gamma}$ when $(1-\tau)^{\frac{1}{\gamma}} \tau$ is also maximized. This would impose a natural borrowing limit for the government.

■

From now we assume that $\bar{b}$ represents the natural borrowing limit. We begin with some useful lemmas

let $L \equiv l^{1+\gamma}$, to make this problem convex,

Substitute for $c(s)$

$$V(b_{-}) = \max_{L(s),b(s)} \sum_{s \in S} \pi(s) \left[\frac{1}{1+\gamma} L(s) + \frac{1}{\beta} P(s) b_{-} - b(s) + \beta V(b(s)) \right]$$

s.t.

$$\begin{aligned} \frac{1}{\beta} P(s) b - b(s) + g(s) &\leq \theta L^{\frac{1}{1+\gamma}}(s) - L(s) \\ b(s) &\leq \bar{b} \\ L(s) &\geq 0. \end{aligned}$$

Lemma 4 *$V(b)$ is stictly concave, continuous, differentiable and $V(b) < \beta^{-1}$ for all $b < \bar{b}$. The feasibility constraint binds for all $b \in (-\infty, \bar{b}]$, $s \in S$ and $(L^*(s))^{1-\frac{1}{1+\gamma}} \geq \frac{1}{1+\gamma}$.*[25]

Proof. *Concavity*

$V(b)$ is concave because we maximize linear objective function over convex set.

Binding feasibility

Suppose that feasibility does not bind for some b, s. Then the optimal $L(s)$ solve $\max_{L(s)\geq 0} \pi(s) \frac{\gamma}{1+\gamma} L(s)$ which sets $L(s) = \infty$. This violates feasility for any finite $b, b(s)$.

Bounds on L

Let $\phi(s) > 0$ be a Lagrange multiplier on the feasibility. The FOC for $L(s)$ is

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

This gives

$$\frac{1}{1+\gamma} L^{\frac{1}{1+\gamma}-1} - \theta = -\frac{1}{\lambda}\frac{\gamma}{1+\gamma} < 0$$

or

$$L^{1-\frac{1}{1+\gamma}} \geq \frac{\theta}{1+\gamma}.$$

Continuity

For any L that satisfy $L^{1-\overline{1+\gamma}} \geq \frac{\theta}{1+\gamma}$, define function Ψ that satisfies $\Psi\left(L^{\frac{1}{1+\gamma}} - \theta L\right) = L$. Since $L^{\frac{1}{1+\gamma}} - L$ is strictly decreasing in L for $L^{1-\frac{1}{1+\gamma}} \geq \frac{1}{1+\gamma}$, this function is well defined. Note that $\Psi()\underbrace{\left(\frac{1}{1+\gamma} L^{\frac{1}{1+\gamma}-1} - \theta\right)}_{<0} = 1$ (so that $\Psi > 0$, i.e. Ψ is strictly decreasing) and $\Psi'' \underbrace{\left(\frac{1}{1+\gamma} L^{\frac{1}{1+\gamma}-1} - 1\right)^2}_{>0} + \underbrace{\Psi}_{<0} \underbrace{\frac{1}{1+\gamma}\frac{\gamma}{1+\gamma} L^{\frac{1}{1+\gamma}-2}}_{<0} = 0$ (so that $\Psi'' \geq 0$, $\Psi'' > 0$, i.e. Ψ is strictly concave on the interior). Ψ is also continuous. When $L^{1-\frac{1}{1+\gamma}} = \frac{1}{1+\gamma}$, $L = (1+\gamma)^{-\frac{(1+\gamma)}{(\gamma)}}$. Let $D \equiv (1+\gamma)^{\frac{-1}{\gamma} - (1+\gamma)^{-\frac{1+\gamma}{(\gamma)}}}$. Then the objective is

$$V(b_{\_}) = \max_{b(s)} \sum_{s\in S} \pi(s)\left[\Psi\left(\frac{1}{\beta}P(s)\,b - b(s) + g(s)\right) + \frac{1}{\beta}P(s)\,b_{\_} - b(s) + \beta V(b(s))\right]$$

s.t.

$$\begin{aligned} b(s) &\leq \bar{b} \\ \frac{1}{\beta}P(s)\,b_{\_} - b(s) + g(s) &\leq D. \end{aligned}$$

[25] This last condition simply means that we do not tax to the right of the peak of the Laffer curve. The revenue maximizing tax is $1 - \bar{\tau} = \frac{1}{1+\gamma}$. At the same time $1 - \tau = l^\gamma$ so if taxes are always to the left of the peak, $\frac{1}{1+\gamma} \leq l^\gamma = \left(L^{\frac{1}{1+\gamma}}\right)^\gamma = L^{1-\frac{1}{1+\gamma}}$.

This function is continuous so V is also continuous.

Differentiability

Continuity and convexity implies differentiability everywhere, including the boundaries.

Strict concavity

Ψ is strictly concave, so on the interior V is strictly concave.

■

Next we characterize policy functions

Lemma 5 *$b(s, b_{-})$ is an increasing function of b for all s for all s, b_{-} where $b(s)$ is interior.*

Proof. Take the FOCs for $b(s)$ from the condition in the previous problem. If $b(s)$ is interior

$$\Psi\left(\frac{1}{\beta}P(s)\,b_{-} - b(s) + g(s)\right) = \beta V(b(s)).$$

Suppose $b_1 < b_2$ but $b_2(s) < b_1(s)$. Then from stict concavity

$$\begin{aligned} V(b_2(s)) &< V'(b_1(s)) \\ \Psi\left(\frac{1}{\beta}P(s)\,b_2 - b_2(s) + g(s)\right) &> \Psi\left(\frac{1}{\beta}P(s)\,b_1 - b_1(s) + g(s)\right). \end{aligned}$$

■

Lemma 6 *There exists an invariant distribution of the stochastic process $b_{t+1} = b(s_{t+1}, b_t)$*

Proof. The state spaces for b_t and s_t are compact. Further the transition function on $s_{t+1}|s_t$ is trivially increasing under i.i.d shocks. We can apply standard arguments as in **?**(see corollary 3) to argue that there exists invariant distribution of assets. ■

Now we characterize the support of this distribution using further properties of the policy rules for $b(s|b_{-})$

Lemma 7 *For any $b_{-} \in (\underline{b}, \bar{b})$, there are s, s'' s.t. $b(s) \geq b_{-} \geq b(s'')$. Moreover, if there are any states s'', s''' s.t. $b(s'') \neq b(s''')$, those inequalities are strict.*

Proof. The FOCs together with the envelope theorem imply that $\mathbb{E}P(s)V'(b(s)) = V'(b_{-}) + \kappa(s)$ We can rewrite this as $\tilde{\mathbb{E}}V'(b(s)) = b + \kappa(s)$ with $\tilde{\pi}(s) = P(s)\pi(s)$

Now if there is at least one $b(s)$ s.t. $b(s) > b_{-}$, by strict concavity of V there must be some s'' s.t. $b(s'') < b$.

If there is at least one $b(s)$ s.t. $b(s) < b_{-}$, the inequality above is strictly only if $b(s''') = \bar{b}$ for some s'''. But $V(\bar{b}) < V(b)$ so there must be some s'' s.t. $b(s'') > b$. Equality is possible only if $b_{-} = b(s)$ for all s. ■

Lemma 8 *Let $\mu(b,s)$ be the optimal policy function for the Lagrange multiplier $\mu(s)$. If $P(s') > P(s'')$ then there exists a $b^*_{s',s''}$ such that for all $b < (>)\ b_{1,s',s''}$ we have $\mu(b,s') > (<)\ \mu(b,s'')$. If $\underline{b} < b^*_{s',s''} < \bar{b}$ then $\mu(b^*_{s',s''},s') = \mu(b^*_{s',s''},s'')$.*

Proof. Suppose that $\mu(b,s') \leq \mu(b,s'')$. Subtracting the implementability for s'' from the implementability constraint for s' we have

$$\begin{aligned}\frac{P(s')-P(s'')}{\beta}b &= S_{s'}(\mu(b,s')) - S_{s''}(\mu(b,s'')) + b'(b,s') - b'(b,s'') \\ &\geq S_{s'}(\mu(b,s')) - S_{s''}(\mu(b,s')) + b'(b,s') - b'(b,s'') \\ &\geq S_{s'}(\mu(b,s')) - S_{s''}(\mu(b,s')) = g(s'') - g(s')\end{aligned}$$

We get the first inequality from noting that $S_s(\mu') \geq S_s(\mu'')$ if $\mu' \leq \mu''$. We obtain the second inequality by noting that $\mu(b,s') \leq \mu(b,s'')$ implies $b'(b,s') \geq b'(b,s'')$ (which comes directly from the concavity of V). Thus, $\mu(b,s') \leq \mu(b,s'')$ implies that

$$b \geq \frac{\beta(g(s'')-g(s'))}{P(s')-P(s'')} = b^*_{s',s''} \tag{46}$$

The converse of this statement is that if $b < b^*_{s',s''}$ then $\mu(b,s') > \mu(b,s'')$. The reverse statement that $\mu(b,s') \geq \mu(b,s'')$ implies $b \leq b^*_{s,s'}$ follows by symmetry. Again, the converse implies that if $b > b^*_{s',s''}$ then $\mu(b,s') < \mu(b,s'')$. Finally, if $\underline{b} < b^*_{s',s''} < \bar{b}$ then continuity of the policy functions implies that there must exist a root of $\mu(b,s') - \mu(b,s'')$ and that root can only be at $b^*_{s',s''}$. ■

Lemma 9 *$P \in \mathcal{P}^*$ is necessary and sufficient for existence of b^* such that $b(s,b^*) = b^*$ for all ss*

Proof. The necessary part follows from taking differences of the (45a) for s',s''. We have

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

Thus $P \in \mathcal{P}^*$. The sufficient part follows from the Lemma 8. If $P \notin \mathcal{P}^*$, equation (46) that defines $b^*_{s',s''}$ will not be same across all pairs. Thus b^* that satisfies $b(s;b^*)$ independent of s will not exist. ■

Lemma 9 implies that under the hypothesis of part 1 of the Theorem 3 there cannot exist an interior absorbing point for the dynamics of debt. This allows us to construct a sequences $\{b_t\}_t$ such that $b_t < b_{t+1}$ with the property that $\lim_t b_t = \underline{b}$. Thus, for any $\epsilon > 0$, there exists a finite history of shocks that can take us arbitrarily close to $\underline{b}$. Since the shocks are i.i.d this finite sequence will repeat i.o. With a symmetric argument we can show that b_t will come arbitrarily close to its upper limit i.o too

Part 2 of Theorem **3**

In this first section we will show that there exists b_1, and if $P(s)$ is sufficiently volatile a b_2, such that if $b_t \leq b_1$ then

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

and if $b_t \geq b_2$ then

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

Recall that b is decreasing in μ, so this implies that if b_t is low (large) enough then there will exist a drift away from the lower (upper) limit of government debt.

With Lemma 8 we can order the policy functions $\mu(b, \cdot)$ for particular regions of the state space. Take b_1 to be

$$b_1 = \min \left\{ b^*_{s',s''} \right\}$$

and WLOG choose $\underline{b} < b_1$. For all $b < b_1$ we have shown that $P(s) > P(s')$ implies that $\mu(b, s) > \mu(b, s')$. The FOC for the problem imply,

$$\mu_{-} = \mathbb{E}P(s)\mu(s) + \underline{\kappa}(s) \tag{47}$$

The inequality in the resource constraint implies that $\phi(s) \geq 0$ implying that $\mu(s) \leq 1$. With some minor algebra algebra we obtain

By decomposing $\mathbb{E}\mu(s)P(s)$ in equation (47), we obtain (using $\mathbb{E}P(s) = 1$)

$$\mu_{-} = \mathbb{E}\mu(s) + \text{cov}(\mu(s), P(s)) + \underline{\kappa}(s) \tag{48}$$

Our analysis has just shown that for $b_{-} < b_1$ we have $\text{cov}_t(\mu(s), P(s)) > 0$ so

$$\mu_{-} > \mathbb{E}\mu(s).$$

If p is sufficiently volatile:

$$P(s') - P(s'') > \frac{\beta(g(s'') - g(s'))}{\bar{b}}$$

then

$$b_2 = \max \left\{ b^*_{s',s''} \right\} < \bar{b}$$

and through a similar argument we can conclude that $\text{cov}(\mu(s), P(s)) < 0$

$$\mu_{-} < \mathbb{E}\mu(s)$$

for $b_- > b_2$ (note $b_- > \underline{b}$ implies $\underline{\kappa}(s) = 0$) which gives us a drift away from the upper-bound.

Part 3 of Theorem **3**

When $P \in \mathcal{P}^*$, Lemma 9 implies existence of b^* as the steady state debt level.

Lemma 10 *There exists μ^* such that μ_t is a sub-martingale bounded above in the region $(-\infty, \mu^*)$ and super-martingale bounded below in the region $(\mu^*, \frac{1}{1+\gamma})$*

Proof. Let μ^* be the associated multiplier, i.e $V_b(b^*) = \mu^*$. Using the results of the previous section, we have that $b_1 = b_2 = b^*$, implying that $\mu_t < (>)\mathbb{E}_t\mu_{t+1}$ for $b_t < (>)b^*$. ■

Lastly we show that $\lim_t \mu_t = \mu^*$. Suppose $b_t < b^*$, we know that $\mu_t > \mu^*$. The previous lemma implies that in this region, μ_t is a super martingale. The lemma 5 shows that $b(s, b_-)$ is continuous and increasing. This translates into $\mu(\mu(b_-), s)$ to be continuous and increasing as well. Thus

$$\mu_t > \mu^* \implies \mu(\mu_t, s_{t+1}) > \mu(\mu^*, s_{t+1})$$

or

$$\mu_{t+1} > \mu^*$$

Thus $\mu*$ provides a lower bound to this super martingale. Using standard martingale convergence theorem converges. The uniqueness of steady state implies that it can only converge to μ^*. For $\mu < \mu^*$, the argument is symmetric.

■

A.5 Proof of Theorem 4

Working with the first order conditions of problem 44, we obtain

$$l(s)^{\gamma} = \frac{\mu(s) - 1}{(1+\gamma)\mu(s) - 1} = 1 - \tau(\mu(s)),$$

implying the relationship between tax rate τ and multiplier μ given by

$$\tau(\mu) = \frac{\gamma\mu}{(1+\gamma)\mu - 1} \tag{49}$$

At the interior, the rest of the first order conditions and the implementability constraints are summarized below

$$\frac{b_{-}P(s)}{\beta} = S(\mu(s), s) + b(s)$$
$$\mu(b_{-}) = \mathbb{E}P(s)\mu(s)$$

where $S(\mu, s)$ is the government surplus in state s given by

$$S(\mu, s) = (1 - \tau(\mu))^{\frac{1}{\gamma}}\tau(\mu) - g(s) = I(\mu) - g(s)$$

The proof of the theorem will have four steps:

Step 1: Obtaining a recursive representation of the optimal allocation in the incomplete markets economy with payoffs $P(s)$ with state variable μ_{-}

Given a pair $\{P(s), g(s)\}$, since $V'(b)$ is one-to-one, so we can re-characterize these equations as searching for a function $b(\mu)$ and $\mu(s|\mu_{-})$such that the following two equations can be solved for all μ_{-}.

$$\frac{b(\mu_{-})P(s)}{\beta} = I(\mu(s)) - g(s) + b(\mu(s)) \tag{50}$$
$$\mu_{-} = \mathbb{E}\mu(s)P(s) \tag{51}$$

Step 2: Describe how the policy rules are approximated

Usually perturbation approaches to solve equilibrium conditions as above look for the solutions to $\{\mu(s|\mu_{-})\}$ and $b(\mu_{-})$ around deterministic steady state of the model. Thus for any b^{ss}, there exists a μ^{ss} that will solve

$$\frac{b^{SS}}{\beta} = I(\mu^{SS}) - \bar{g} + b^{SS}$$

For example if we set the perturbation parameter q to scale the shocks, $g(s) = \mathbb{E}g(s)+q\Delta_g(s)$ and $P(s) = 1 + q\Delta_P(s)$, the first order expansion of $\mu(s|\mu_-)$ will imply that it is a martingale. Such approximations are not informative about the ergodic distribution. [26]

In contrast we will approximate the functions $\mu(s|\mu_-)$ around around economy with payoffs in $\bar{P} \in \mathcal{P}^*$.

In contrast we a) explicitly recognize that policy rules depend on payoffs: $\mu(s|\mu_-, \{P(s)\}_s)$ and $b(\mu_-, \{P(s)\}_s)$ and then take take the first order expansion with respect to both μ_- and $\{P(s)\}$ around the vector $(\bar{\mu}, \{\bar{P}(s)\}_s)$ where $\bar{P}(s) \in \mathcal{P}^*$: these payoffs support an allocation such that limiting distribution of debt is degenerate around the some value $\bar{b}$; and $\overline{\mu}$ is the corresponding limiting value of multiplier. The next two expression make the link between $\bar{\mu}$ and $\bar{b}$ explicit.

$$\bar{b} = \frac{\beta}{1-\beta}\left(I(\overline{\mu}) - \overline{g}\right) \tag{52a}$$

where $\overline{g} = \mathbb{E}g$ and $\overline{p}$ as

$$\overline{P}(s) = 1 - \frac{\beta}{\bar{b}}(g(s) - \overline{g}) \tag{52b}$$

As noted before $b(\overline{\mu}; \overline{p}) = \bar{b}$ solves the the system of equations (50-51) for $\mu'(s) = \overline{\mu}$ when the payoffs are $\overline{P}(s)$

We next obtain the expressions that characterize the linear approximation of $\mu(s|\mu_-, \{P(s)\})$ and $(\mu_-, \{P(s)\})$ around some arbitrary point $(\bar{\mu}, \{\bar{P}(s)\}_s)$ where $\bar{P}(s) \in \mathcal{P}^*$. We will use these expressions to compute the mean and variance of the ergodic distribution associated with the approximated policy rules. Finally as a last step we propose a particular choice of the point of approximation.

The derivatives $\frac{\delta\mu(s|\mu_-,\{P(s)\}}{\delta\mu_-}, \frac{\delta\mu(s|\mu_-,\{P(s)\}}{\delta P(s)}$ and similarly for $b(\mu_-, \{P(s)\}$ are obtained below:

Differentiating equation (50) with respect to μ around $(\overline{\mu}, \overline{P})$ we obtain

$$\frac{\overline{P}(s)}{\beta}\frac{\partial b}{\partial \mu_-} = \left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu_-}\right]\frac{\partial \mu(s)}{\partial \mu_-}.$$

Differentiating equation (51) with respect to μ_- we obtain

$$1 = \sum_s \pi(s)\overline{P}(s)\frac{\partial \mu'(s)}{\partial \mu_-}$$

combining these two equations we see that

$$\frac{1}{\beta}\left(\sum_s \pi(s)\overline{P}(s)^2\right)\frac{\partial b}{\partial \mu_-} = I'(\overline{\mu}) + \frac{\partial b}{\partial \mu_-}$$

[26] One can do higher order approximations, but part 3 of theorem 3 hints that for economies with payoffs close to $\mathcal{P}^*$, the stochastic steady state in general is far away from μ^{SS}.

Noting that $\mathbb{E}\overline{P}^2(s) = 1 + \frac{\beta^2}{\overline{b}^2}\sigma_g^2$ we obtain

$$\frac{\partial b}{\partial \mu_-} = \frac{I'(\overline{\mu})}{\frac{\beta}{\overline{b}^2}\sigma_g^2 + \frac{1-\beta}{\beta}} < 0 \tag{53}$$

as $I'(\overline{\mu}) < 0$. We then have directly that

$$\frac{\partial \mu'(s)}{\partial \mu} = \frac{\overline{P}(s)}{\frac{\beta^2}{\overline{b}^2}\sigma_g^2 + 1} = \frac{\overline{P}(s)}{\mathbb{E}\overline{P}(s)^2} \tag{54}$$

We can perform the same procedure for $P(s)$. Differentiating equation (50) with respect to $P(s)$ we around $(\overline{\mu}, \overline{p})$ we obtain

$$\frac{\overline{p}(s')}{\beta}\frac{\partial b}{\partial P(s)} + 1_{s,s'}\frac{\overline{b}}{\beta} - \frac{\pi(s)\overline{b}\overline{p}(s')}{\beta} = \left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]\frac{\partial \mu(s')}{\partial P(s)} \tag{55}$$

Here $1_{s,s'}$ is 1 if $s = s'$ and zero otherwise. Differentiating equation (51) with respect to $P(s)$ we obtain

$$0 = \pi(s)\overline{\mu} - \pi(s)\overline{\mu} + \sum_{s'}\pi(s)\overline{p}(s')\frac{\partial \mu(s')}{\partial P(s)} = \sum_{s'}\pi(s')\overline{p}(s')\frac{\partial \mu(s')}{\partial P(s)}$$

Again we can combine these two equations to give us

$$\frac{\mathbb{E}\overline{p}(s)^2}{\beta}\frac{\partial b}{\partial P(s)} + \frac{\pi(s)\overline{b}}{\beta}(\overline{p}(s) - \mathbb{E}\overline{p}(s)^2) = 0$$

or

$$\frac{\partial b}{\partial P(s)} = \pi(s)\overline{b}\frac{\mathbb{E}\overline{p}^2 - \overline{p}(s)}{\mathbb{E}\overline{p}^2} \tag{56}$$

Going back to equation (55) we have

$$\frac{\partial \mu(s')}{\partial P(s)} = \frac{\overline{b}}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]}\left(1_{s,s'} - \frac{\pi(s)\overline{p}(s)\overline{p}(s')}{\mathbb{E}\overline{p}^2}\right) \tag{57}$$

Step 3: Getting expressions for the mean and variance of the ergodic distribution around some arbitrary point of approximation

For an arbitrary $\left(\overline{\mu}, \{\overline{P}(s)\}_s\right)$, using the derivatives that we computed, we can characterize the dynamics of $\hat{\mu} \equiv \mu_t - \overline{\mu}$ using our approximated policies.

$$\hat{\mu}_{t+1} = B(s_{t+1})\hat{\mu}_t + C(s_{t+1}),$$

where $B(s)$ and $C(s)$ are respective derivatives. Note that both are random variables and let us denote their means $\overline{B}$ and $\overline{C}$, and variances σ_B^2 and σ_C^2 . Suppose that $\hat{\mu}$ is distributed according to the ergodic distribution of this linear system with mean $\mathbb{E}\hat{\mu}$ and variance σ_μ^2. Since

$$B\hat{\mu} + C,$$

has the same distribution we can compute the mean of this distribution as

$$\begin{aligned}\mathbb{E}\hat{\mu} &= \mathbb{E}\left[B\hat{\mu}+C\right]\\ &= \mathbb{E}\left[\mathbb{E}_{\hat{\mu}}\left[B\hat{\mu}+C\right]\right]\\ &= \mathbb{E}\left[\overline{B}\hat{\mu}+\overline{C}\right]\\ &= \overline{B}\mathbb{E}\hat{\mu}+\overline{C}\end{aligned}$$

solving for $\mathbb{E}\hat{\mu}$ we get

$$\mathbb{E}\hat{\mu} = \frac{\overline{C}}{1-\overline{B}} \tag{58}$$

For the variance $\sigma^2_{\hat{\mu}}$ we know that

$$\sigma^2_{\hat{\mu}} = \mathrm{var}(B\hat{\mu}+C) = \mathrm{var}(B\hat{\mu}) + \sigma^2_C + 2\mathrm{cov}(B\hat{\mu}, C)$$

Computing the variance of $B\hat{\mu}$ we have

$$\begin{aligned}\mathrm{var}(B\hat{\mu}) &= \mathbb{E}\left[(B\hat{\mu}-\overline{B}\mathbb{E}\hat{\mu})^2\right]\\ &= \mathbb{E}\left[(B\hat{\mu}-\overline{B}\hat{\mu}+\overline{B}\hat{\mu}-\overline{B}\mathbb{E}\hat{\mu})^2\right]\\ &= \mathbb{E}\left[\mathbb{E}_{\hat{\mu}}\left[(B-\overline{B})^2\hat{\mu}^2+2(B-\overline{B})(\hat{\mu}-\mathbb{E}\hat{\mu})\overline{B}\mathbb{E}\hat{\mu}+(\hat{\mu}-\mathbb{E}\hat{\mu})^2\bar{B}^2\right]\right]\\ &= \mathbb{E}\left[\sigma^2_B\hat{\mu}^2+(\hat{\mu}-\mathbb{E}\hat{\mu})^2\overline{B}\right]\\ &= \sigma^2_B(\sigma^2_{\hat{\mu}}+(\mathbb{E}\hat{\mu})^2)+\sigma^2_{\hat{\mu}}\overline{B}^2\end{aligned}$$

while for the covariance of $B\hat{\mu}$ and C

$$\mathrm{cov}(B\hat{\mu}, C) = \sigma_{BC}\mathbb{E}\hat{\mu}$$

Putting this all together we have

$$\sigma^2_{\hat{\mu}} = \frac{\sigma^2_B(\mathbb{E}\hat{\mu})^2+\sigma_{BC}\mathbb{E}\hat{\mu}+\sigma^2_C}{1-\overline{B}^2-\sigma^2_B} \tag{59}$$

Step 4: Choice of the point of approximation

To get the expressions in Theorem 3, we finally choose a particular $\overline{P} = P^*(s) \in \mathcal{P}^*$. This will be the closest complete market economy to our the given $P(s)$ in L^2 sense. Formally,

$$\min_{\tilde{P}\in\mathcal{P}^*} \sum_s \pi(s)(P(s)-\tilde{P}(s))^2.$$

Since all payoffs in $\mathcal{P}^*$ are associated with some b^* and μ^* via equations (52), we can re write the above problem as choosing $\overline{\mu}$ so as to minimize the variance of the difference between $P(s)$

and the set of steady state payoffs. Let P^* be the solution to this minimization problem. The first order condition for this linearization gives us

$$2\sum_{s'} \pi(P(s') - P^*(s', \mu^*))\frac{\delta P^*(s, \mu^*)}{\delta \mu^*} = 0$$

as noted before

$$P^*(s) = 1 - \frac{\beta}{b^*(\mu^*)}\left(g(s) - \mathbb{E}g\right)$$

thus

$$\frac{\delta P^*}{\delta \mu^*} \propto P^* - 1$$

Thus we can see the the optimal choice of $\overline{\mu}$ is equivalent to choosing $\overline{\mu}$ such that

$$\begin{aligned}
0 &= \sum_{s'} \pi(s')(P(s') - P^*(s', \mu^*))(P^*(s', \mu^*) - 1) \\
&= -\sum_{s'} \pi(s')(P(s') - P^*(s', \mu^*)) + \sum_{s'} \pi(s')(P(s') - P^*(s', \mu^*))P^*(s', \mu^*) \\
&= \sum_{s'} \pi(s')(P(s') - P^*(s', \mu^*))P^*(s', \mu^*) \\
&= \mathbb{E}\left[(P - P^*)P^*\right]
\end{aligned} \tag{60}$$

At these values of $\overline{p} = P^*$ and $\overline{\mu} = \mu^*$ we have that C for our linearized system is

$$C(s') = \sum_s \left\{ \frac{b^*}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]}\left(1_{s,s'} - \frac{\pi(s)P^*(s)P^*(s')}{\mathbb{E}\overline{p}^2}\right)(P(s) - P^*(s)) \right\}$$

Taking expectations we have that

$$\begin{aligned}
\overline{C} &= \sum_s \left\{ \frac{b^*}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]}\left(\pi(s) - \frac{\pi(s)P^*(s)}{\mathbb{E}\overline{p}^2}\right)(P(s) - P^*(s)) \right\} \\
&= \frac{b^*}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]}\left(\mathbb{E}(P - \overline{p}) - \frac{\mathbb{E}\left[(P - \overline{p})\overline{p}\right]}{\mathbb{E}\overline{p}^2}\right) \\
&= 0
\end{aligned} \tag{61}$$

Thus the linearized system will have the same mean for μ, $\overline{\mu}$, as the closest approximating steady state payoff structure.

We can also compute the variance of the ergodic distribution for μ. Note

$$
\begin{aligned}
C(s') &= \sum_s \left\{ \frac{b^*}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]} \left(1_{s,s'} - \frac{\pi(s)P^*(s)P^*(s')}{\mathbb{E}P^{*2}}\right)(P(s) - P^*(s)) \right\} \\
&= \frac{b^*}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]} \left(P(s') - P^*(s') - P^*(s')\frac{\sum_s \pi(s)P^*(s)(p_s - P^*(s))}{\mathbb{E}P^{*2}}\right) \\
&= \frac{b^*}{\beta\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]}(P(s') - P^*(s))
\end{aligned}
$$

As noted before

$$
\sigma_\mu^2 = \frac{b^{*2}}{\beta^2\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]^2\left(1 - \overline{B}^2 - \sigma_B^2\right)}\|P - P^*\|^2
$$

The variance of government debt in the linearized system is

$$
\sigma_b^2 = \frac{b^{*2}\left(\frac{\partial b}{\partial \mu}\right)^2}{\beta^2\left[I'(\overline{\mu}) + \frac{\partial b}{\partial \mu}\right]^2\left(1 - \overline{B}^2 - \sigma_B^2\right)}\|P - P^*\|^2
$$

This can be simplified using the following expressions:

$$
I'(\overline{\mu}) + \frac{\partial b}{\partial \mu} = \frac{\mathbb{E}P^{*2}}{\beta}\frac{\partial b}{\partial \mu},
$$

$$
\overline{B} = \frac{1}{\mathbb{E}P^{*2}}
$$

and

$$
\sigma_B^2 = \frac{\mathrm{var}(P^*)}{(\mathbb{E}P^{*2})^2}
$$

to

$$
\sigma_b^2 = \frac{b^{*2}}{\mathbb{E}P^{*2}\mathrm{var}(P^*)}\|P - P^*\|^2 \tag{62}
$$

Noting that $\mathbb{E}P^{*2} = 1 + \mathrm{var}(P^*) > 1$, we have immediately that up to first order the relative spread of debt is bounded by

$$
\frac{\sigma_b}{b^*} \leq \sqrt{\frac{\|P - P^*\|^2}{\mathrm{var}(P^*)}} \tag{63}
$$

A.6 Proof of Theorem 5

Proof. Using Theorem 1 let $\tilde{b} = b_1 - b_2$. Under the normalization that $b_2 = 0$, the variable $\tilde{b}$ represents public debt government or the assets of the productive agent. The optimal plan solves the following Bellman equation,

$$V(\tilde{b}_{\text{-}}) = \max_{c_1(s),c_2(s),b'(s)} \sum_s \pi(s) \left\{ \omega \left[c_1(s) - \frac{l_1^{1+\gamma}(s)}{1+\gamma} \right] + (1-\omega)c_2(s) + \beta V(\tilde{b}(s)) \right\} \tag{64}$$

subject to

$$c_1(s) - c_2(s) + \tilde{b}(s) = l(s)^{1+\gamma} + \beta^{-1} P(s) \tilde{b}_{\text{-}} \tag{65a}$$

$$nc_1(s) + (1-n)c_2(s) + g(s) \leq \theta_2 l(s) n \tag{65b}$$

$$c_2(s) \geq 0 \tag{65c}$$

$$\overline{b} \geq \tilde{b}(s) \geq \underline{b} \tag{65d}$$

Let $\mu(s), \phi(s), \lambda(s), \underline{\kappa}(s), \overline{\kappa}(s)$ be the Lagrange multipliers on the respective constraints. The FOC are summarized below

$$\omega - \mu(s) = n\phi(s) \tag{66a}$$

$$1 - \omega + \mu(s) - \phi(s)(1-n) + \lambda(s) = 0 \tag{66b}$$

$$-\omega l^{\gamma}(s) + \mu(s)(1+\gamma)l^{\gamma}(s) + n\phi(s)\theta = 0 \tag{66c}$$

$$\beta V'(\tilde{b}(s)) - \mu(s) - \overline{\kappa}(s) + \underline{\kappa}(s) = 0 \tag{66d}$$

and the envelope condition

$$V'(\tilde{b}_{\text{-}}) = \sum_s \pi(s)\mu(s)\beta^{-1}P(s) \tag{66e}$$

■

To show part 1 of Theorem 5, we show that $\frac{\omega}{n} > \frac{1+\gamma}{\gamma}$ is sufficient for the Lagrange multiplier $\lambda(s)$ on the non-negativity constraint to bind.

Lemma 11 *The multiplier on the budget constraint $\mu(s)$ is bounded above*

$$\mu(s) \leq \min\left\{\omega - n, \frac{\omega}{1+\gamma}\right\}$$

Similiarly the multiplier of the resource constraint is bounded below,

$$\phi(s) \geq \max\left\{1, \frac{\omega}{n}\left[\frac{\gamma}{1+\gamma}\right]\right\}$$

Proof.

Notice that the labor choice of the productive household implies $\frac{1}{1-\tau} = \frac{\theta_2}{l^{\gamma}(s)}$.

As taxes go to $-\infty$ (66c) implies that $\mu(s)$ approaches $\frac{\omega}{1+\gamma}$ from below. Similiarly the non-negativity of $c_2(s)$ imposes a lower bound of 1 on $\phi(s)$. This translates into an upper bound of $\omega - n$ on μ. ■

Lemma 12 *There exists a $\bar{\omega}$ such that $\omega > \bar{\omega}$ implies $c_2(s) = 0$ for all b*

Proof.

By the KKT conditions $c_2(s) = 0$ if $\lambda(s) > 0$. Now (66b) implies this is true if $\mu(s) < \omega - n$. The previous lemma bounds $\mu(s)$ by $\frac{\omega}{1+\gamma}$.

We can thus define $\bar{\omega} = n\left(\frac{1+\gamma}{\gamma}\right)$ as the required threshold Pareto weight to ensure that the unproductive agent has zero consumption forever.

■

Now for the rest of the parts $\omega < n\left(\frac{1+\gamma}{\gamma}\right)$, we can have positive transfers for low enough public debt. In particular, we can define a maximum level of debt $\mathcal{B}$ that is consistent with an interior solution for the unproductive agents' consumption.

Guess an interior solution $c_{2,t} > 0$ or $\lambda_t = 0$ for all t. This gives us $l_t = l^*$ defined below:

$$l^* = \left[\frac{n\theta}{\omega - (\omega - n)(1+\gamma)}\right]^{\frac{1}{\gamma}} \tag{67}$$

As long as $\omega < n\left(\frac{1+\gamma}{\gamma}\right)$ At the interior solution $\tilde{b}(s) = \tilde{b}_{-}$ and using the implementability constraint and resource constraints (65a) and (65b) respectively, we can obtain the expression for $c_2(s)$

$$c_2(s) = n\theta l^* - n{l^*}^{1+\gamma} - \tilde{b}_{-}(1 - P(s)\beta^{-1}) - g(s)$$

Non-negativity of c_2 implies,

$$\tilde{b}_{-} \leq \frac{g(s) - n\theta l^* + n{l^*}^{1+\gamma}}{\beta^{-1}P(s) - 1}$$

We can also express this as

$$\tilde{b}_- \leq \frac{g(s) - \tau^* y^*}{\beta^{-1} P(s) - 1},$$

where the right hand side of the previous equation is just the present discounted value of the primary deficit of the government at the constant taxes τ^* associated with l^* defined in (67). As long as $\beta^{-1}P(s) - 1 > 0$, this object is well defined, we define $\mathcal{B} = \min_s \left[\frac{g(s) - n\theta l^* + n l^{*1+\gamma}}{\beta^{-1} P(s) - 1}\right]$. Thus for $\tilde{b}_- < \mathcal{B}$ the optimal allocation has constant taxes given by τ^* and debt $\tilde{b}_-$, while transfers are given by

$$T(s) = n\theta l^* - n l^{*1+\gamma} - \tilde{b}_-(1 - P(s)\beta^{-1}) - g(s),$$

and are strictly positive.

In the next lemma we show how $\mathcal{B}$ varies with ω.

Lemma 13 *For $\omega \leq n\frac{1+\gamma}{\gamma}$, we have $\frac{\partial \mathcal{B}}{\partial \omega} > 0$.*

Proof. The sign of the derivative of $\mathcal{B}$ with respect to ω is the same as the sign of the following derivative:

$$\frac{\partial \left[l^{*1+\gamma} - \theta l^*\right]}{\partial \omega}$$

Note that (67) implies that l^* is increasing in ω. Note that,

$$\frac{\partial \left[l^{*1+\gamma} - \theta l^*\right]}{\partial \omega} = \frac{\partial l^*}{\partial \omega}\left[(1+\gamma) l^{*\gamma} - \theta\right]$$

So the sign of the required derivative depends on $[(1+\gamma)l^{*\gamma} - \theta]$. We now argue that this expression is positive over the range $\omega \leq n\frac{1+\gamma}{\gamma}$.

Again from the expression for l^*, we see that

$$\min_{\omega \leq n\frac{1+\gamma}{\gamma}} l^{*\gamma} = \frac{\theta}{1+\gamma}$$

.

Thus we can see that $\mathcal{B}$ is increasing in ω

■

For initial debt greater than $\mathcal{B}$, we distinguish cases when payoffs are perfectly aligned with $g(s)$ i.e belong to the set $\mathcal{P}^*$ and when they are not. For part 2 case b, let $P \notin \mathcal{P}^*$.

Lemma 14 *There exists a $\check{b} > \mathcal{B}$ such that there are two shocks $\underline{s}$ and $\overline{s}$ and the optimal choice of debt starting from $\tilde{b}_- \leq \check{b}$ satisfies the following two inequalities:*

$$\tilde{b}(\underline{s}, \tilde{b}_-) > \mathcal{B}$$

$$\tilde{b}(\overline{s}, \tilde{b}_-) \leq \mathcal{B}$$

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

$$\check{b} = \mathcal{B} + \frac{\epsilon\beta}{2P(\overline{s})}$$

Now suppose to the contrary $\tilde{b}(\overline{s}, \tilde{b}_{\text{-}}) > \mathcal{B}$ for some $\tilde{b}_{\text{-}} \leq \check{b}$. This implies that $\tau(s, \tilde{b}_{\text{-}}) > \tau^*$ and $T(\overline{s}, \tilde{b}_{\text{-}}) = 0$.

The government budget constraint implies

$$\frac{P(\overline{s})\tilde{b}_{\text{-}}}{\beta} + g(s) = \tilde{b}(\overline{s}, \tilde{b}_{\text{-}}) + (1 - \tau(\overline{s}, \tilde{b})_{\text{-}})l(\overline{s}, \tilde{b}_{\text{-}}).$$

As,

$$\frac{P(\overline{s})\tilde{b}_{\text{-}}}{\beta} + g(\overline{s}) \leq \frac{P(\overline{s})\mathcal{B}}{\beta} + g(\overline{s}) + \frac{\epsilon}{2} < \frac{P(\overline{s})\mathcal{B}}{\beta} + g(\overline{s}) + \epsilon$$

This further implies,

$$\tilde{b}(\overline{s}, \tilde{b}_{\text{-}}) + (1 - \tau(\overline{s}, \tilde{b}_{\text{-}}))l(\tau(\overline{s}, \tilde{b}_{\text{-}})) > [\tilde{b}(\overline{s}, \tilde{b}_{\text{-}}) + (1 - \tau^*)l^* > \mathcal{B} + (1 - \tau^*)l^* > \frac{P(\overline{s})\tilde{b}_{\text{-}}}{\beta} + g(\overline{s}) + T(\overline{s}, \tilde{b}_{\text{-}}) = \frac{P(\overline{s})\tilde{b}_{\text{-}}}{\beta} + g(\overline{s}) + \epsilon.$$

Combining the previous two inequalities yields a contradiction. The other inequality, $\tilde{b}(\underline{s}, \tilde{b}_{\text{-}}) > \mathcal{B}$ follows from the definition of $\mathcal{B}$. This is because if it was not true then $\tilde{b}(s, \tilde{b}_{\text{-}}) \leq \mathcal{B}$ for all shocks. This implies that the solution is interior. However the only initial conditions that have this property are less than equal to $\mathcal{B}$.

Now define $\overline{\mu}(\tilde{b}(s, \tilde{b}_{\text{-}}))$ as $\max_s \mu(s, \tilde{b}_{\text{-}})$ and $\hat{s}(\tilde{b}_{\text{-}})$ as the shock that achieves this maximum. Now we show that $\hat{\mu}(\tilde{b}(s, \tilde{b}_{\text{-}}))$ is finite for all $b_{\text{-}} \leq \overline{b}$. We show the claim for the natural debt limit.

Let $b^n(s) = (\beta^{-1}P(s) - 1)^{-1}\left[\theta^{\frac{\gamma}{1+\gamma}}\left(\frac{1}{1+\gamma}\right)^{\frac{1}{\gamma}}\left(\frac{\gamma}{1+\gamma}\right) - g(s)\right]$ be the maximum debt supported by a particular shock s. The natural debt limit is defined as $\overline{b}^n = \min_s b^n(s)$. Note that $\lim_{b \to \overline{b}^n} \mu(\tilde{b}_{\text{-}}) = \infty$

Now choose s such that $b^n(s) > \overline{b}^n$ and consider the debt choice next period for the same shock s when it comes in with debt $\overline{b}^n$.

Suppose it chooses $\tilde{b}(s, \overline{b}^n) = \overline{b}^n$, then taxes will have to be set to $\frac{\gamma}{1+\gamma}$ and the tax income will be $\frac{\gamma}{1+\gamma}l(\frac{\gamma}{1+\gamma}) = \theta^{\frac{\gamma}{1+\gamma}}\left(\frac{1}{1+\gamma}\right)^{\frac{1}{\gamma}}\left(\frac{\gamma}{1+\gamma}\right)$. The budget constraint will then imply that,

$$\frac{\overline{b}^n P(s)}{\beta} + g(s) = \theta^{\frac{\gamma}{1+\gamma}}\left(\frac{1}{1+\gamma}\right)^{\frac{1}{\gamma}}\left(\frac{\gamma}{1+\gamma}\right) + \overline{b}^n$$

$$\overline{b}^n = (P(s)\beta^{-1} - 1)^{-1}\left(\theta^{\frac{\gamma}{1+\gamma}}\left(\frac{1}{1+\gamma}\right)^{\frac{1}{\gamma}}\left(\frac{\gamma}{1+\gamma}\right) - g(s)\right)$$

However the right hand side is the definition of $b^n(s)$ and,

$$b^n(s) > \overline{b}^n.$$

Thus we have a contradiction and the optimal choice of debt at the natural debt limit $\tilde{b}(s, \overline{b}^n) < \overline{b}^n$.

This inturn means that $\lim_{\tilde{b} \to \overline{b}^n} \overline{\mu}(\tilde{b}) < \infty$.

Now note that $\overline{\mu}(\tilde{b}_{\_}) - \mu(\tilde{b}_{\_})$ is continuous on $[\check{b}, \quad \overline{b}^n]$ and is bounded below by zero, therefore attains a minimum at $\tilde{b}^{min}$. Let $\delta = \hat{\mu}(\tilde{b}^{min}) - \mu(\tilde{b}^{min}) > \eta > 0$. If this was not true then $P(s) \in \mathcal{P}^*$ as μ will have an absorbing state.

Let $\mu(\omega, n) = \omega - n$. This is the value of μ when debt falls below $\mathcal{B}$.

Now consider any initial $\tilde{b}_{\_} \in [\mathcal{B}, \overline{b}^n]$. If $\tilde{b}_{\_} \leq \check{b}$, then by lemma 14, we know that $\mathcal{B}$ will be reached in one shock. Otherwise if $\tilde{b}_{\_} > \check{b}$, we can construct a sequence of shocks $s_t = \hat{s}(\tilde{b}_{t-1})$ of length $N = \frac{\mu(\omega,n) - \mu(\tilde{b}_{\_})}{\delta}$. There exits $t < N$ such that $\tilde{b}_t < \check{b}$, otherwise,

$$\mu_t > \mu(\tilde{b}_{\_}) + N\delta > \mu(\omega, n)$$

Thus we can reach $\mathcal{B}$ in finite steps. Since shocks are i.i.d, this is an almost sure statement. At $\mathcal{B}$, transfers are strictly positive for some shocks $T_t > 0$ a.s. and taxes are given by τ^*.

Now consider the payoffs $P \in \mathcal{P}^*$ such that the associated steady state debt $b^* > \mathcal{B}$. Under the guess $T_t = 0$, the same algebra as in Theorem 3 goes through and we can show that $\tilde{b}_{\_} = b^*$ is a steady state for the heterogeneous agent economy. Thus the heterogeneous agent economy for a given $P \in \mathcal{P}^*$ has a continuum of steady states given by the set $[\bar{b}, \quad \mathcal{B}] \cup \{b^*\}$.

In the region $\tilde{b}_{\_} > b^*$, as before μ_t is supermartingale bounded below by b^*. Since there is a unique fixed point in the region $\tilde{b}_{\_} \in [b^*, \overline{b}^n]$, μ_t converges to μ^* associated with b^*. Transfers are zero and taxes are given by τ^{**}

$$\tau^{**} = \frac{\gamma \mu^*}{(1+\gamma)\mu^* - 1} \tag{68}$$

In the region $[\mathcal{B}, \quad b^*]$ the outcomes depend on the exact sequence of shocks we can show that μ_t is a submartingale. This follows from the observation that for all $\tilde{b}_{\_} > \mathcal{B}$, we have $T(s) = 0$ and the outcomes from the representative agent economy allow us to order $\mu(s)$ relative $P(s)$. At $\tilde{b}_{\_} = \mathcal{B}$, $\mu(s) = \omega - n$ and is constant. Thus in the region $[\mathcal{B}, \quad B^*]$, μ_t is sub martingale and it converges. However if $\tilde{b}_t$ gets sufficiently close to $\check{b}$, then it can converge to $\mathcal{B}$ and if it gets sufficiently close to b^*, it can converge to b^*. Either of this can happen with strictly positive probability. ■

A.7 Proof of Theorem 6

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

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

subject to the constraints

$$1 + \rho'(s)[l_1(s)^2 - 1] + \beta x'(s) - \frac{x\frac{P(s)}{c_2(s)}}{\mathbb{E}[\frac{P(s)}{c_2(s)}]} = 0 \tag{69}$$

$$\mathbb{E}\frac{P(s)}{c_1(s)}(\rho'(s) - \rho) = 0 \tag{70}$$

$$\theta_1(s)l_1(s) - c_1(s) - c_2(s) - g = 0 \tag{71}$$

$$\rho'(s)c_2(s) - c_1(s) = 0 \tag{72}$$

where the $\pi(s)$ is the probability distribution of the aggregate state s. If we let $\pi(s)\mu(s)$, λ, $\pi(s)\xi(s)$ and $\pi(s)\phi(s)$ be the Lagrange multipliers for the constraints (69)-(72) respectively then we obtain the following FONC for the planners problem [27]

$c_1(s)$:

$$\frac{\alpha_1\pi(s)}{c_1(s)} - \frac{\lambda\pi(s)}{c_1(s)^2}(\rho'(s) - \rho) - \pi(s)\xi(s) - \pi(s)\phi(s) = 0 \tag{73}$$

$c_2(s)$:

$$\frac{\alpha_2\pi(s)}{c_2(s)} + \frac{xP(s)\pi(s)}{c_2(s)^2\mathbb{E}[\frac{P}{c_2}]}\left[\mu(s) - \frac{\mathbb{E}[\mu\frac{P}{c_2}]}{\mathbb{E}[\frac{P}{c_2}]}\right] - \pi(s)\xi(s) + \pi(s)\rho'(s)\phi(s) = 0 \tag{74}$$

$l_1(s)$:

$$-\alpha_1\pi(s)l_1(s) + 2\mu(s)\pi(s)\rho'(s)l_1(s) + \theta_1(s)\pi(s)\xi(s) = 0 \tag{75}$$

$x'(s)$:

$$V_x(x'(s),\rho'(s)) + \mu(s) = 0 \tag{76}$$

$\rho'(s)$:

$$\beta V_\rho(x'(s),\rho'(s)) + \frac{\lambda\pi(s)}{c_1(s)} + \mu(s)[l_1(s)^2 - 1] + \pi(s)\phi(s)c_2(s) = 0 \tag{77}$$

[27] Appendix A.8 discuses the associated second order conditions that ensure these policies are optimal

In addition there are two envelope conditions given by

$$V_x(x,\rho) = -\sum_{s'} \frac{\mu(s')\pi(s')\frac{P(s)}{c_2(s')}}{\mathbb{E}[\frac{P}{c_2}]} = -\frac{\mathbb{E}[\mu\frac{P}{c_2}]}{\mathbb{E}[\frac{P}{c_2}]} \tag{78}$$

$$V_\rho(x,\rho) = -\lambda\mathbb{E}[\frac{P}{c_1}] \tag{79}$$

In the steady state, we need to solve for 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 (69)-(79) are satisfied when $\rho'(s) = \rho$ and $x'(s) = x$. It should be clear that if we replace $\mu(s) = \mu$, equation (76) and the envelope condition with respect to x is always satisfied. Additionally under this assumption equation (74) simplifies significantly,since

$$\frac{xP(s)\pi(s)}{c_2(s)^2\mathbb{E}[\frac{P}{c_2}]}\left[\mu(s) - \frac{\mathbb{E}[\mu\frac{P}{c_2}]}{\mathbb{E}[\frac{P}{c_2}]}\right] = 0$$

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

$$1 + \rho[l_1(s)^2 - 1] + \beta x - \frac{xP(s)}{c_2(s)\mathbb{E}[\frac{P}{c_2}]} = 0 \tag{80}$$

$$\theta_1(s)l_1(s) - c_1(s) - c_2(s) - g = 0 \tag{81}$$

$$\rho c_2(s) - c_1(s) = 0 \tag{82}$$

$$\frac{\alpha_1}{c_1(s)} - \xi(s) - \phi(s) = 0 \tag{83}$$

$$\frac{\alpha_2}{c_2(s)} - \xi(s) + \rho\phi(s) = 0 \tag{84}$$

$$[2\mu\rho - \alpha_1]l_1(s) + \theta_1(s)\xi(s) = 0 \tag{85}$$

$$\lambda\left(\frac{P(s)}{c_1(s)} - \beta\mathbb{E}\left[\frac{P}{c_1}\right]\right) + \mu[l_1(s)^2 - 1] + \phi(s)c_2(s) = 0 \tag{86}$$

We can rewrite equation (83) 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 (84) and normalizing $\alpha_1 + \alpha_2 = 1$ we obtain

$$\xi(s) = \frac{1}{(1+\rho)\,c_2(s)} \tag{87}$$

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

$$\phi(s) = \frac{\alpha_1 - \rho\alpha_2}{(\rho(1+\rho))\,c_2(s)} \tag{88}$$

From equation (80) we can solve for $l_1(s)^2 - 1$ as

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

This can be used along with equations (86) and (88) to obtain

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

Note that the LHS depends on s while the RHS does not, hence the solution to this equation is

$$\lambda = -\frac{\mu x}{\mathbb{E}[\frac{P}{c_2}]} \tag{89}$$

and

$$\mu = \frac{\alpha_1 - \rho\alpha_2}{1+\rho} \tag{90}$$

Combining these with equation (85) 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 2 *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 (81) 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

$$l_1(s) = \frac{g + \sqrt{g^2 + 4C(\rho)\theta_1(s)^2}}{2\theta_1(s)} \tag{91}$$

and

$$\frac{1}{c_2(s)} = \left(\frac{1+\rho}{C(\rho)}\right)\left(\frac{g + \sqrt{g^2 + 4C(\rho)\theta_1(s)^2}}{2\theta_1(s)^2}\right) \tag{92}$$

A steady state is then a value of ρ such that

$$x(s) = \frac{1 + \rho[l_1(\rho, s)^2 - 1]}{\frac{P(s)/c_2(\rho,s)}{\mathbb{E}[\frac{P}{c_2}](\rho)} - \beta} \tag{93}$$

s independent of s.

The following lemma, which orders consumption and labor across states, will be useful in proving the parts of theorem **??**. 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 15 *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. ■

Now we use these lemma to prove the part 1 and part 2 of theorem 6

Proof.

[Part 1.] For a riskfree bond when $P(s) = 1$. In order for there to exist a ρ such that equation (93) is independent of the state (and hence have a steady state) we need the existence of root for the following function

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

From lemma 15 we can conclude that

$$1+\rho[l_1(\rho,s_l)^2-1] > 1+\rho[l_1(\rho,s_h)^2-1] \tag{94}$$

and

$$\frac{1/c_2(\rho,s_l)}{\mathbb{E}[\frac{P}{c_2}](\rho)}-\beta > \frac{1/c_2(\rho,s_h)}{\mathbb{E}[\frac{P}{c_2}](\rho)}-\beta \tag{95}$$

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[28]

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

We can also show that

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

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

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

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

Thus, there exists $\overline{\rho}$ such that $1+\overline{\rho}[l_1(\overline{\rho},s_l)^2-1]=0$. [29] From equation (94), 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

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

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

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

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

[Part 2] 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 $P(s)$ such that $\beta<\frac{P(s)/c_2(\rho_{SS},s)}{\mathbb{E}[\frac{P}{c_2}]}$ such that

$$1>\frac{1+\rho_{SS}[l_1(\rho_{SS},s_l)^2-1]}{1+\rho_{SS}[l_1(\rho_{SS},s_h)^2-1]}=\frac{\frac{P(s_l)/c_2(\rho_{SS},s_l)}{\mathbb{E}[\frac{P}{c_2}]}-\beta}{\frac{P(s_h)/c_2(\rho_{SS},s_h)}{\mathbb{E}[\frac{P}{c_2}]}-\beta}\tag{96}$$

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

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

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

Thus, in the steady state, the government is holding debt, under the normalization that the unproductive worker holds no assets. Note this imposes a restriction of $\frac{P(s_l)}{P(s_h)}$.

$$\frac{P(s_l)c_2^{-1}(\rho_{SS}, s_l) - \beta \mathbb{E} P c_2^{-1}}{P(s_h)c_2^{-1}(\rho_{SS}, s_h) - \beta \mathbb{E} P c_2^{-1}} < 1$$

or

$$\frac{P(s_l)}{P(s_h)} < \frac{c_2^{-1}(\rho_{SS}, s_h)}{c_2^{-1}(\rho_{SS}, s_l)} < 1$$

or

Thus $P(s_l) < P(s_h)$ i.e payoffs have to be sufficiently procyclical.

■

A.8 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.

$$V(\boldsymbol{x}, \boldsymbol{\rho}) = \max_{c_i(s), l_i(s), \boldsymbol{x}'(s), \boldsymbol{\rho}'(s)} \sum_s P(s) \left(\left[\sum_i \pi_i \alpha_i U(c_i(s), l_i(s)) \right] + \beta(s) V(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s)) \right) \tag{97}$$

$$U_{c,i}(s)c_i(s) + U_{l,i}(s)l_i(s) - \rho_i'(s)\left[U_{c,1}(s)c_1(s) + U_{l,1}(s)l_1(s)\right] + \beta(s)x_i'(s) = \frac{x_i U_{c,i}(s)}{\mathbb{E} U_{c,i}} \tag{98a}$$

$$\sum_s \Pr(s) U_{c,1}(s)(\rho_i(s) - \rho_i) = 0 \tag{98b}$$

$$\frac{\rho'(s)}{\theta_1(s)} U_{l,1}(s) = \frac{1}{\theta_i(s)} U_{l,i}(s) \tag{98c}$$

$$\sum_{j=0}^{I} \pi_j c_j(s) + g(s) = \sum_{j=0}^{I} \pi_j \theta_j(s) l_j(s) \tag{98d}$$

$$U_{c,i}(s) = \rho_i'(s) U_{c,1}(s) \tag{98e}$$

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

$c_1(s)$**:**

$$\pi_1\alpha_1 U_{c,1}(s) + \sum_{i=2}^{I}\left(\mu_i'(s)\rho_i'(s)\right)\left[U_{cc,1}(s)c_1(s) + U_{c,1}(s)\right] + \lambda U_{cc,1}(s)\sum_{i=2}^{I}(\rho_i'(s) - \rho_i) - \pi_1\xi(s) + \sum_{i=2}^{N}\zeta_i(s)\rho_i'(s)U_{cc,1}(s) = 0 \tag{99a}$$

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

$$\pi_i\alpha_i U_{c,i}(s) - \mu_i'(s)\left[U_{cc,i}(s)c_i(s) + U_{c,i}(s)\right] + \frac{x_i U_{cc,i}(s)}{\mathbb{E}U_{c,i}}\left(\mu_i'(s) - \frac{\mathbb{E}\mu_i' U_{c,i}}{\mathbb{E}U_{c,i}}\right) - \pi_i\xi(s) - \zeta_i(s)U_{cc,i}(s) = 0 \tag{99b}$$

$l_1(s)$**:**

$$\pi_1\alpha_1 U_{l,1}(s) + \sum_{i=2}^{I}\mu_i'(s)\rho_i(s)\left[U_{ll,1}(s)l_1(s) + U_{l,1}(s)\right] - \sum_{i=2}^{N}\frac{\rho_i'(s)\phi_i(s)}{\theta_1(s)}U_{ll,1}(s) + \pi_1\theta_1(s)\xi(s) = 0 \tag{99c}$$

$l_2(s)$**:**

$$\pi_i\alpha_i U_{l,i}(s) - \mu_i'(s)\left[U_{ll,i}(s)l_i(s) + U_{l,i}(s)\right] + \frac{\phi_i(s)}{\theta_i(s)}U_{ll,i}(s) + \pi_i\theta_i(s)\xi(s) = 0 \tag{99d}$$

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

$$\beta(s)V_{\rho_i}(\boldsymbol{x}'(s), \boldsymbol{\rho}_i'(s)) + \mu_i'(s)\left[U_{c,1}(s)c_1(s) + U_{l,1}(s)l_1(s)\right] + \lambda_i U_{c,1}(s) - \phi_i(s)\frac{U_{l,1}(s)}{\theta_1(s)} + U_{c,1}(s)\zeta_i(s) = 0 \tag{99e}$$

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

$$V_{x_i}(\boldsymbol{x}'(s), \boldsymbol{\rho}'(s)) - \mu_i'(s) = 0. \tag{99f}$$

Equations (98a)-(98e) and (99a)-(99e) then define the necessary conditions for an interior maximization of the planners problem for the state $(\boldsymbol{x}, \boldsymbol{\rho})$. In addition to these we have the two envelop conditions

$$V_{x_i}(\boldsymbol{x}, \boldsymbol{\rho}) = \frac{\sum_s P(s)\mu_i'(s)U_{c,i}(s)}{\mathbb{E}U_{c,i}(s)} = \frac{\mathbb{E}\mu_i' U_{c,i}}{\mathbb{E}U_{c,i}}, \tag{100a}$$

and

$$V_{\rho_i}(\boldsymbol{x}, \boldsymbol{\rho}) = -\lambda_i \mathbb{E}U_{c,1}. \tag{100b}$$

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 $\boldsymbol{\mu}$ must follow the weighted martingale

$$\mu_i - \frac{\sum_s P(s)\mu_i'(s)U_{c,i}(s)}{\sum_s P(s)U_{c,i}(s)} = 0. \tag{101}$$

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 (98a)-(99e)and (101), 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 $\overline{z} = z(\overline{y})$

$$D_zF(\overline{z}, \overline{y}, g(\overline{z}))D_yz(\overline{y}) + D_yF(\overline{z}, \overline{y}, g(\overline{z})) + D_gF(\overline{z}, \overline{y}, g(\overline{z}))Dg(\overline{z})D_yz(\overline{z}) = 0$$

In order to linearize $z(y)$ around the steady state $\overline{y}$ we need to compute $D_yz(\overline{y})$. The envelope condition (100b) 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

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

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

$$D_zF(\overline{z}, \overline{y}, \overline{w})D_yz(\overline{y}) + D_yF(\overline{z}, \overline{y}, \overline{w}) + D_wF(\overline{z}, \overline{y}, \overline{w}) \begin{pmatrix} Dw(\overline{z})D_yz(\overline{y}) & 0 \\ 0 & Dw(\overline{z})D_yz(\overline{y}) \end{pmatrix} \Phi D_yz(\overline{z}) = 0, \tag{102}$$

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 (29). 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 (26) at the steady state. We can then apply second order tests to verify that the first order necessary conditions are sufficient.

A.9 Proof for Theorem 7

Proof.

The state at time t can be written as

$$\hat{\Psi}_t = C_t C_{t-1} \cdots C_1 \hat{\Psi}_0.$$

where the C_i are all random variables being $C(s)$ with probability $\pi(s)$. Taking expectations and applying independence we then obtain

$$
\begin{aligned}
\mathbb{E}_0[\hat{\Psi}_t] &= \mathbb{E}_0[C_t C_{t-1} \cdots C_1]\hat{\Psi}_0 && (103)\\
&= \mathbb{E}[C_t]\mathbb{E}[C_{t-1}] \cdot \mathbb{E}[C_1]\hat{\Psi}_0 && (104)\\
&= \overline{C}^t \hat{\Psi}_0 && (105)
\end{aligned}
$$

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

$$
\bar{\hat{\Psi}}_{t|0} \equiv \mathbb{E}_0[\hat{\Psi}_t] = \overline{C}^t \hat{\Psi}_0 \to \mathbf{0}. \tag{106}
$$

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

$$
\Sigma_{\Psi,t|0} = \mathbb{E}_0\left[(\hat{\Psi}_t - \bar{\hat{\Psi}}_t)(\hat{\Psi}_{t|0} - \bar{\hat{\Psi}}_{t|0})'\right]
$$

or

$$
\Sigma_{\Psi,t|0} = \mathbb{E}_0 \hat{\Psi}_t \hat{\Psi}_t' - \bar{\hat{\Psi}}_{t|0} \bar{\hat{\Psi}}_{t|0}'. \tag{107}
$$

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

$$
\Sigma_{\Psi,t|0} \approx \mathbb{E}[C\Sigma_{\Psi,t-1|0}C'.] \tag{108}
$$

Showing that $\hat{\Psi}_{t|0} \to \mathbf{0}$ in distribution, amounts to showing that $\Sigma_{\Psi,t|0} \to 0$ for any starting point Σ_Ψ and following the process in equation (108). One can obtain a necessary condition for $\|\Sigma_{\Psi,t|0}\| \to 0$ under the process in equation (108). That process can be rewritten as follows

$$
\begin{aligned}
\Sigma_{\Psi,t|0} &= \mathbb{E}[C\Sigma_{\Psi,t-1|0}C'] && (109)\\
&= \sum_s \Pr(s) C(s) \Sigma_{\Psi,t-1|0} C(s)' && (110)\\
&= \sum_s \Pr(s)(\overline{C} + (C(s) - \overline{C}))\Sigma_{\Psi,t-1|0}(\overline{C} + (C(s) - \overline{C}))' && (111)\\
&= \overline{C}\Sigma_{\Psi,t-1|0}\overline{C}' + \sum_s \Pr(s)(C(s) - \overline{B})\Sigma_{\Psi,t-1|0}(C(s) - \overline{C})'. && (112)
\end{aligned}
$$

This is a deterministic linear system in $\Sigma_{\Psi,t|0}$. Suppose we reshape $\Sigma_{\Psi,t|0}$ as a vector (denoted by $\text{vec}(\Sigma_{\Psi,t|0})$) and let $\hat{C}$ be a (square) matrix such that equation 112 is written as

$$
\text{vec}(\Sigma_{\Psi,t|0}) = \hat{C}\text{vec}(\Sigma_{\Psi,t-1|0}).
$$

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

We used theorem 7 to verify local stability of a wide range of examples. The typical finding is that the steady state is generically stable and that convergence is slow. In figure 8 we plot the comparative statics for the dominant eigenvalue and the associated half-life for a two-agent economy with CES preferences. We set the other parameter to match a Frisch elasticity of 0.5, a real interest rate of 2%, marginal tax rates around 20%, and a 90-10 percentile ratio of wage earnings of 4. In the first exercise, we vary the size of the expenditure shock keeping risk aversion σ at one The x- axis plots the spread in expenditure normalized by the undistorted GDP and reported in percentages. In the bottom panel, we fix the size of shock such that it produces a 5% fall in expenditure fall at risk aversion of one and vary σ from 0.8 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 reasonable values of curvatures and shocks. Both increasing the size of the shock or risk aversion increases the volatility of the interest rates, speeding up the transition towards the steady state.

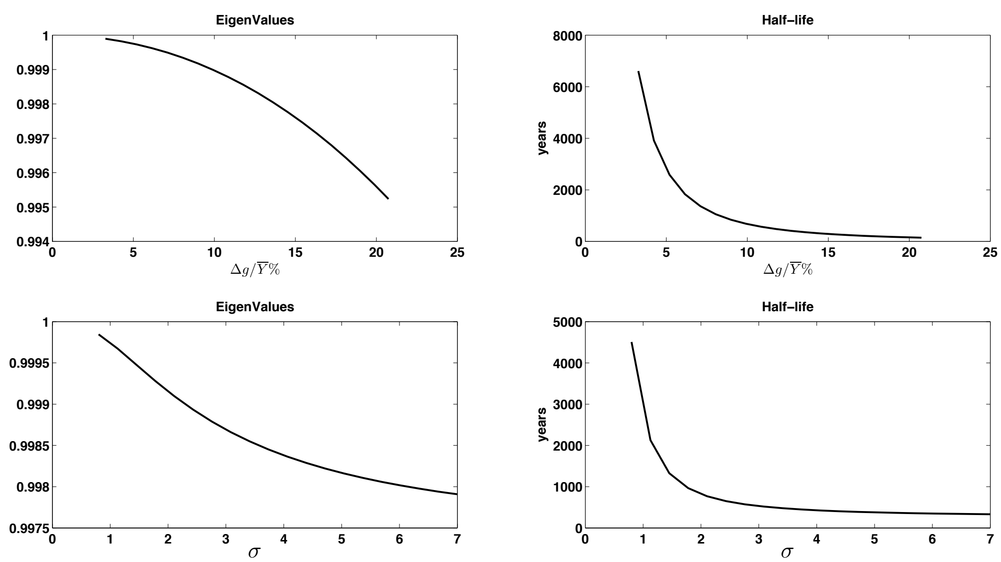


Figure 8: The top (bottom) panel plots the dominant eigenvalue of $\hat{C}$ and the associated half life as we increase the spread between the expenditure levels (risk aversion).

A.10 Numerical approximation to a Ramsey plan

This appendix applies a method of Evans (2014) to approximate a Ramsey plan for the I type of agents economy of section 7. To focus on essential steps, we take a special case where productivities for agent i are $\log \theta_{i,t} = \log \overline{\theta}_i + \sigma_\epsilon \epsilon_t$ and where ϵ_t is i.i.d with mean zero and variance one.[30] We will use the following notation for the rmainder of this appendix:

- Vectors $z_{i,t-1} \in \mathcal{R}^{n_z}$ of state variables for agents of type $i = 1, \ldots, I$.
- Vectors $y_{i,t} \in \mathcal{R}^{n_y}$ of choice variables for agents of type $i = 1, \ldots, N$; the vector $y_{i,t}$ often includes components of $z_{i,t}$.
- A vector $Y_t \in \mathcal{R}^{n_Y}$ of policy variables chosen by a Ramsey planner
- A distribution Γ_{t-1} over $z_{i,t-1} \in \mathcal{R}^{n_z}$. This can be a measure confined to I mass points. For a I type case with fixed masses of agents and no idiosyncratic risk, Γ_{t-1} has all information that is in the set $\{z_{i,t-1}\}_i$. However the algorithm applies for more general economies with idiosyncratic risk, and there the notation Γ_{t-1} helps.

A *Ramsey plan* can be represented as a set of functions (y, Y, Γ) defining the recursions:

$$\begin{aligned} y_{i,t} &= y_i(\epsilon_t; z_{i,t-1}, \Gamma_{t-1}, \sigma_\epsilon), \quad i = 1, \ldots I \\ Y_t &= Y(\epsilon_t; \Gamma_{t-1}, \sigma_\epsilon) \\ \Gamma_t &= \Gamma(\epsilon_t; \Gamma_{t-1}) \end{aligned} \tag{113}$$

These functions appropriately organize solutions to the implementability and optimality conditions associated with a Ramsey plan, conditions that can be expressed in the following forms:

$$\mathbb{E}_{t-1} F(y_{i,t}, \mathbb{E}_{t-1} y_{i,t}, Y_t, y_{i,t+1}, \epsilon_t; z_{i,t-1}, \sigma_\epsilon) = 0, \tag{114}$$

which must hold for all $z_{i,t-1}$ such that $\Gamma_{t-1}(z_{i,t-1}) > 0$; and

$$\int G(y_{i,t}, Y, \epsilon_t; z_{i,t-1}, \sigma_\epsilon) d\Gamma_{t-1} = 0. \tag{115}$$

The terms $E_{t-1} y_{i,t}$ in equation (114) capture constraints requiring that subsets of individual decision variables $y_{i,t}$ must be measurable with respect to time $t-1$ information. Our goal is to approximate outcomes generated by the system of equations (113). To do this, we generate a sequence of approximations to the system of equations (113). We generate the t^{th} outcome

[30] Extending the method to more general productivity process as in (31a) is straightforward.

along a sample path by drawing pseudo random vectors ϵ_t and applying our approximation of equations (113) for date t. Our approximation to these functions at date t depends on the outcomes $\{z_{i,t-1}\}_i, \Gamma_{t-1}$ generated at the previous step of the simulation. To generate functions approximating (113) at date t, we use a small-noise expansion (i.e., around $\sigma_\epsilon = 0$) to those functions at state $\{z_{i,t-1}\}_i, \Gamma_{t-1}$, an expansion that exploits economic properties associated with the limiting $\sigma_\epsilon = 0$ economy at that state. Thus, to approximate sample paths drawn from the recursive system (113), we use a sequence of Taylor series approximations around a sequence of points generated endogenously during a simulation.

Anmol, David, and Tom: we need a sentence or two right here describing the virtues of the approximation.

The steps of the algorithm proceed sequentially as follows:

1. Given some Γ_{t-1}, compute the individual and aggregate choice variables in a limiting economy with $\sigma_\epsilon = 0$. For our problem, we choose state variables that ensure that $\Gamma_t = \Gamma_{t-1}$ in this limiting economy.[31] The allocation in the limiting economy is a set of values $\left(\{\bar{y}_i,\}_i, \bar{Y}\right)$ that solve(114) and (115) at $\sigma_\epsilon = 0$. This logic gives us a set of non-linear equations

$$F(\bar{y}_i, \bar{y}_i, \bar{Y}, \bar{y}_i, 0; \bar{z}_{i,t-1}, 0) = 0 \quad \forall i \tag{116}$$

$$\int G(\bar{y}, \bar{Y}, 0; z_{i,t-1}, \sigma_\epsilon) d\Gamma_{t-1} = 0 \tag{117}$$

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

2. Next construct a truncated Taylor series approximation to the functions y, Y appearing in (113) around the steady state $\bar{y}(\Gamma_{t-1})$ and $\bar{Y}(\Gamma_{t-1})$ obtained in the step 1; this yields approximations

$$\begin{aligned} y(\epsilon_t; z_{i,t-1}, \Gamma_{i,t-1}, \sigma_\epsilon) \approx \bar{y}(\Gamma_{t-1}) &+ \frac{\partial y}{\partial \epsilon}(0; z_{i,t-1}, \Gamma_{t-1}, 0)\epsilon_t \\ &+ \frac{1}{2}\frac{\partial^2 y}{\partial \epsilon^2}(0; z_{i,t-1}, \Gamma_{t-1}, 0)\epsilon_t^2 \\ &+ \frac{1}{2}\frac{\partial^2 y}{\partial \sigma_\epsilon^2}(0; z_{i,t-1}, \Gamma_{t-1}, 0)\sigma_\epsilon^2 \end{aligned} \tag{118}$$

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

and

$$
\begin{aligned}
Y(\epsilon_t;\Gamma_{t-1},\sigma_\epsilon) \approx \bar{Y}(\Gamma_{t-1}) &+ \frac{\partial Y}{\partial \epsilon}(0;\Gamma_{t-1},0)\epsilon_t \\
&+ \frac{1}{2}\frac{\partial^2 Y}{\partial \epsilon_t^2}(0;\Gamma_{t-1},0)\epsilon_t^2 \\
&+ \frac{1}{2}\frac{\partial^2 Y}{\partial \sigma_\epsilon^2}(0;\Gamma_{t-1},0)\sigma_\epsilon^2.
\end{aligned}
\tag{119}
$$

The main computational task is to evaluate derivatives at the steady state. This involves totally differentiating system (114) and (115) at the non stochastic steady state associated with Γ_{t-1}. [32]

3. Draw shocks ϵ_t and use the approximate policies in (118) and (119) to obtain $y_{i,t}$ and Y_t. Remember that $z_{i,t}$ is assumed to be included in the vector $y_{i,t}$. This yields us the next Γ_t.

4. Advance to $t+1$ and repeat steps 1 to 3 using the updated Γ_t as the initial distribution.

A key feature of this algorithm is how the points of approximation, and hence the derivatives that capture how agents respond to aggregate shocks, vary along a history. This feature is particularly attractive for problems where the mean of the ergodic distribution can be sufficiently far away from the initial conditions and the convergence to the ergodic distribution is slow. Since ours is a perturbation approach, handling a high dimensional $z_{i,t-1}$ is more tractable than projection methods using finite order polynomials as basis functions, as in Judd et al. (2011)

Application to our problem

We describe what (y_i, Y, z_i) and conditions (114) and (115) are for our problem.

- Individual states $z_i = (m_i, \mu_i)$ and the implied distribution Γ over $z = (m, \mu) \in \mathbb{R}^2$ with I mass points,
- individual's choice variables $y_i = (m_i, \mu_i, c_i, l_i, \phi_i, x_{i\text{-}}, \rho_i, , \psi_{i\text{-}})$, and
- planner's aggregate choice variables $Y = (\tau_l, T, \alpha, \xi, \gamma_{-})$,

The objects listed above are defined now using modifications of the Bellman equation for $t \geq 1$ that we spelled out in problem (22) in section 6.1. For convenience in applying our algorithm, we rewrote the Bellman equation. Let $x_{\text{-}} = U^i_{c\text{-}} b_i$ and $m^i{}_{\text{-}} \propto \frac{1}{U^i_c}$ with $\sum_i m^i = 1$.

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

Note that Ricardian equivalence implies that we can normalize $\sum_i \frac{x^i_-}{U^i_{c-}} = 0$. Thus, the dimension of the state variables appearing in our modified Bellman equation is also $2I-2$, as in problem (22) in section 4. However, by not normalizing with respect to some arbitrary agent's asset holdings (such as $i=1$), we attain a symmetry that turns out to be convenient.[33]

The modified Bellman equation for $t \geq 1$ is:

$$V(\vec{x}_-, \vec{m}_-) = \max \sum_s \left[\Pi(s) \sum_i \omega^i U(c^i(s), l^i(s)) + \beta V(\vec{x}(s), \vec{m}(s)) \right] \tag{120}$$

subject to

$$\begin{aligned}
\mu^i(s): \quad & \frac{P(s) U^i_c(s) x^i_-}{\beta \mathbb{E}_- P U^i_c} = U^i_c(s)(c^i(s) - T(s)) + U^i_l(s) l^i(s) + x^i(s) & (121)\\
\phi^i(s): \quad & U^i_c(s) \exp \epsilon(s) \theta^i (1 - \tau_l(s)) = -U^i_l(s) & (122)\\
\rho^i(s): \quad & \alpha(s) = m^i(s) U^i_c(s) & (123)\\
\psi^i_-: \quad & \gamma_- = m^i_- \mathbb{E}_- P U^i_c & (124)\\
\xi(s): \quad & \sum_i n^i \left[\exp \epsilon(s) \theta^i l^i(s) - c^i(s) \right] = 0 & (125)\\
& \sum_i \frac{x^i(s)}{m^i(s)} = 0 & (126)\\
& \sum_i m^i(s) = 1 & (127)
\end{aligned}$$

Satisfying equation (124) amounts to imposing the definition of $m^i(s)$, while equation (125) imposes that all agents face the same price of the asset. In particular, the existence of a γ_- that satisfies (125) implies that for any (i,j) pair

$$\frac{m^i{}_-}{m^j{}_-} = \frac{\mathbb{E} P U^j_c}{P \mathbb{E} U^i_c}.$$

Using the definition of m^i, we get

$$\frac{U^i_c(s_-)}{U^j_c(s_-)} = \frac{\mathbb{E} P U^i_c}{\mathbb{E} P U^j_c}.$$

Thus, we see that equations (124) and (125) correspond to equations (23e) and (23b) of the formulation in the 6.1. (F, G) in (114) and (115) become the FOCs of problem (120) and equations (122) - (127). Solutions to (F, G) in our model have multiple recursive representations.

In particular, we have alternatives that express allocations as:

[33] This also makes it natural to extend the algorithm to the case with a continuum of agents.

- functions of $(\vec{x}_-, \vec{m}_-, s)$ and $\vec{\mu}_-(\vec{x}_-, \vec{m}_-)$ with endogenous law of motion for $(\vec{x}_-, \vec{m}_-)$

  or

- functions of $(\vec{\mu}_-, \vec{m}_-, s)$ and $\vec{x}_-(\vec{\mu}_-, \vec{m}_-)$ with endogenous law of motion for $(\vec{\mu}_-, \vec{m}_-)$

We employ the second alternative because when $\sigma_\epsilon = 0$, for $I > 2$ that there are multiple μ_- associated with an arbitrary x_- although only one survives in the limit $\sigma_\epsilon \to 0$. However, making μ_- to be the state retrieves a unique x_-.[34]

We include productivity processes like equation (31a) by adding an auxiliary aggregate state variable $\mathcal{Q}_t$ that is a vector of length N and that stores the quantiles of productivities of agents $1, 2, \ldots, I$, respectively. Its (trivial) law of motion is $\mathcal{Q}_t = \mathcal{Q}_{t-1}$. This allows us to express the vector of wages as a function of the shock and the state variable $\mathcal{Q}_t$.

References

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

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.

[34] We solved in both ways (using projection methods) for $I = 2$ to verify that they produce same allocation.

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.

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.

Evans, David. 2014. "Perturbation Theory with Heterogeneous Agents."

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. 2014. "The Nature of Countercyclical Income Risk." *Journal of Political Economy*, 122(3): pp. 621–660, URL: `http://www.jstor.org/stable/10.1086/675535`.

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`.

Kuhn, Moritz. 2014. "Trends in income and wealth inequality."

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, 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`.

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

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

Quadrini, Vincenzo, and Jose-Victor Rios-Rull. 2014. "Inequality in Macroeconomics."

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`.

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): .

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. 2012. “Optimal Fiscal Policy in an Economy with Private Borrowing Limits.” DOI: `http://dx.doi.org/10.1111/jeea.12010`.

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`.