Fiscal Policy and Debt Management with Incomplete Markets*

Anmol Bhandari
bhandari@umn.edu

David Evans
devans@uoregon.edu

Mikhail Golosov
golosov@princeton.edu

Thomas J. Sargent
thomas.sargent@nyu.edu

June 28, 2016

Abstract

A Ramsey planner chooses a distorting tax on labor and manages a portfolio of securities in an economy with incomplete markets. We develop a method that uses second order approximations of Ramsey policies to obtain formulas for conditional and unconditional moments of government debt and taxes that include means and variances of the invariant distribution as well as speeds of mean reversion. We establish that asymptotically the planner's portfolio minimizes a measure of fiscal risk. Analytic expressions that approximate moments of the invariant distribution apply to data on a primary government deficit, aggregate consumption, and returns on traded securities. For U.S. data, we find that an optimal target debt level is negative but close to zero, that the invariant distribution of debt is very dispersed, and that mean reversion is slow.

Key words: Distorting taxes. Spanning. Transfers. Optimal Portfolio. Government debt.

1 Introduction

This paper models a Ramsey planner who optimally manages a portfolio of debts and other securities to smooth fluctuations in tax distortions in an incomplete markets economy hit by aggregate shocks. Within a production economy without capital, the government raises revenue by issuing securities and imposing linear taxes on labor income, then spends on exogenous government expenditures, payouts on government securities, and transfers. The government and private agents trade an exogenously specified set of risky securities whose returns depend on the aggregate state. An economy with complete markets and an economy with a one-period risk-free bond only are interesting special cases.

We make extensive use of an approximation to a Ramsey plan that we construct from second order perturbations around current levels of government debt. We confirm that these quadratic

*We thank Fernando Alvarez, David Backus, V.V. Chari, Maryam Farboodi, Emmanuel Farhi, Xavier Gabaix, Lars Peter Hansen, Ali Shourideh, Pierre Yared, Stephen Zeldes and seminar participants at numerous conferences and seminars for helpful comments

approximations are accurate by comparing them to the solutions obtained using numerical methods. Under conditions that we describe, the approximating laws of motion are linear functions of the aggregate shocks and the current level of government debt. Our quadratic approximations then enable analytic and interpretable expressions for means, variances, and rates of convergence to an invariant distribution of debt, tax revenues, and tax rates.[1] Empirical counterparts to our expressions for these objects can be constructed from data on the primary government deficit, aggregate consumption, and returns on securities traded by the government. We show that asymptotically the government's optimal debt portfolio minimizes a particular criterion measuring fiscal risk.

To isolate underlying principles, we start with a baseline setting in which agents have quasi-linear preferences and the market structure is restricted to a single security whose payout we allow to be correlated with the government purchase process. The joint distribution of returns and government purchases is i.i.d. over time. From the planner's Euler equations, we establish existence of an invariant distribution of government debt. Up to third-order terms, we show that the drift in the dynamics of debt is proportional to the covariance of returns with total government spending (debt service plus exogenous government purchases). A level of debt that minimizes the variance of total government spending sets this covariance to zero and serves as a point of attraction for the stochastic process for debt. The speed of mean reversion is inversely proportional to the variance of the return on the security, and the variance of the invariant distribution is proportional to the amount of risk that the government bears at its risk-minimizing debt level. Later sections of the paper show that the principle that government debt approaches a level that minimizes fiscal risk extends well beyond our baseline case.

Allowing trade in more securities yields additional insights. If returns satisfy a spanning condition, the planner can replicate a complete markets allocation like Lucas and Stokey's (1983). When that spanning condition is not satisfied, being able to trade more securities decreases the speed of convergence to the invariant distribution because additional securities facilitate hedging and thereby lower the cost of being away from a long-run target level of government debt. By assuming two particular securities, a consol and a short-term security, we derive prescriptions for optimal maturity management. In this two-security case, the riskiness of the return on the short-maturity asset relative to that on the consol affects the average maturity of the total debt. In particular, if the return on the long maturity bond is riskier than the return on the short-maturity bill, then the optimal maturity of the planner's portfolio is inversely proportional to total public debt and most adjustment to aggregate shocks is done with the bill. We extend the analysis to incorporate risk aversion and more general shock processes. We show that insights from the baseline model apply provided that we use concepts of "effective returns" and "effective shocks" – returns on the government debt portfolio and innovations to the present discounted value of the primary government deficit adjusted by marginal utilities of consumption, respectively.

In a quantitative section, we seek two goals: (i) to compare two numerical approximations, one using a global method, the other using formulas derived from our quadratic approximation; and (ii) to study the predictions of the model for realistic shock and return processes. To this end, we use U.S. data to calibrate plausible shock and return processes. Our analytical

[1] We can also use our quadratic approximation to get analytic expressions for other moments.

expressions derived under simpler environments continue to be accurate in the calibrated model. We find that the optimal level of debt is close to zero and that the optimal policy displays slow mean reversion (a half-life of 250 years). These results are driven by the fact that a significant amount of variation in returns to the U.S. portfolio is uncorrelated with output; that means that holding large quantities of debt or assets would frustrate hedging objectives.

To focus on some important forces, our paper obviously shuts down forces emphasized in other theories of optimal levels of government debt. For example, by allowing a government each period to choose whether or not to service its debt, the literature on sovereign debt focuses attention on how the adverse consequences of default endogenously generates incentives to repay debt obligations. The government in our model has no default option and requires no such incentives. This eliminates the design of incentives to induce payment as determinants of the level of government debt and its maturity composition and puts the hedging considerations on which we focus front and center. Our model describes optimal fiscal policy of a government that never contemplates dishonoring its debts. (We like to think of the U.S. and some European governments as being in this situation.) Additionally, we focus on real debt. Extending our approach to economies with possibilities of default and monetary economies is straightforward but space-consuming as it would require us to introduce several layers of additional complications to our model. We leave that for future work.

1.1 Relationships to literatures

Our paper builds on a large literature about a Ramsey planner who chooses a competitive equilibrium with distorting taxes once-and-for-all at time zero.[2] Many of these papers assume either complete markets as in Lucas and Stokey (1983), Buera and Nicolini (2004), Angeletos (2002), or a singe-period risk-free bond only and quasilinear preferences as in Barro (1979) and Aiyagari et al. (2002). In contrast, our analysis allows a more general incomplete markets structure and risk-aversion. In both complete market economies and quasilinear settings with a risk-free bond only, *any* level of debt is optimal in the sense that the Ramsey planner sets a time 0 conditional mathematical expectation of public debt in all future periods equal to initial debt. We show that this result is not generic: small departures from the assumptions in those earlier papers imply that, driven by hedging considerations, starting from any initial debt, government debt converges to a unique risk-minimizing level.

In a related context, Barro (1999) and Barro (2003) study tax smoothing in an environment in which revenue needs are deterministic but refinancing opportunities are stochastic. In Barro's setting, it is optimal for a government to issue a consol as a way to insulate inter-temporal tax smoothing motives from concerns about rolling over short maturity debt at uncertain prices. In contrast, our analysis allows both revenue needs and returns on the debt to be stochastic. We estimate empirically relevant properties of returns on debt and then find an optimal government portfolio associated with those returns.

Technically, our paper is closely related to Aiyagari et al. (2002). Those authors include an analysis of an economy in which a representative agent has quasilinear preferences. In addition to a linear labor tax, they allow a uniform nonnegative lump sum transfer. There is a continuum of

[2] For instance Lucas and Stokey (1983), Aiyagari et al. (2002), Chari et al. (1994), Farhi (2010).

invariant distributions for debt, all of which feature a zero labor tax rate and debt levels that are negative and sufficiently large in absolute value to finance all government expenditures from the government's interest revenues, with nonnegative transfers absorbing all aggregate fluctuations by adjusting one-to-one with the aggregate shock. A key difference in our paper is how we treat transfers. While Aiyagari et al. (2002) exogenously restrict transfers, we deduce optimal transfers from an explicit redistribution motive by including agents who sometimes cannot afford to pay positive lump sum taxes. We show that so long as the utility functions of those agents are strictly concave and the planner cares about them, the result of Aiyagari et al. (2002) about properties of the invariant distributions goes away. A benevolent planner in such settings wants to minimize fluctuations of both tax rates *and* transfers; he ultimately targets a (generally unique) level of debt that minimizes risk. The invariant distribution studied by Aiyagari et al. (2002) emerge only in a limit as the risk-aversion of all recipients of transfers goes to zero.

The equilibrium approximation tools that we apply in this paper are complementary to ones used by Faraglia et al. (2012), Lustig et al. (2008), and Siu (2004), who numerically study optimal Ramsey plans in specific incomplete markets settings. Our approximation method allows us to derive closed form expressions for the invariant distribution of debt and taxes that illuminate underlying forces. Our work is also related to Debortoli et al. (2016 forthcoming) who numerically characterize optimal debt management when a government cannot commit to future taxes.

Our theory of government portfolio management shares features of the single-investor optimal portfolio theory of Markowitz (1952) and Merton (1969). Bohn (1990) and Lucas and Zeldes (2009) use insights from the single-investor literature to study portfolio choices of a government in partial equilibrium settings after having specified a government loss function. Common to both Merton's investor and to our Ramsey planner are hedging motives that shape portfolio and savings choices. However, unlike Merton's investor, our Ramsey planner is benevolent (it maximizes the utility of the agents with whom it trades) and it internalizes the general equilibrium effects of its distorting tax rate choices on equilibrium prices and quantities. As a consequence, the optimal portfolio strives to minimize a measure of fiscal risk and is not preoccupied with the usual mean-variance trade-off.

The remainder of this paper is organized as follows. In Section 2, we analyze a streamlined setting in which only one risky security can be traded and the representative agent has quasilinear preferences. In Section 3, we extend the analysis to include multiple assets, persistent shocks, concerns for redistribution, and risk aversion. In Section 4, we study a quantitative example with parameters calibrated to U.S. data.

2 Quasilinear preferences

We begin with a streamlined setting. Time is discrete and infinite with periods denoted $t = 0, 1, ...$ Each of a measure one of identical agents has preferences over consumption and labor supply sequences $\{c_t, l_t\}_t$ that are ordered by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(c_t - \frac{1}{1+\gamma} l_t^{1+\gamma} \right), \tag{1}$$

where $\mathbb{E}_t$ is a mathematical expectations operator conditioned on time t information and $\beta \in (0,1)$ is a time discount factor. One unit of labor produces one unit of a nonstorable single good that can be consumed by households or the government. Feasibility requires

$$c_t + g_t = l_t, \quad t \geq 0, \tag{2}$$

where g_t denotes government consumption.

The government imposes a flat tax at rate τ_t on labor earnings and buys or sells a single one-period security having an exogenous state-contingent payoff p_t. Consumers sell or buy that same security, so it is in zero net supply each period. Let B_t be the number of securities that the government sells in period t at price q_t. Government budget constraints are

$$g_t + p_t \mathrm{B}_{t-1} = \tau_t l_t + q_t \mathrm{B}_t, \quad t \geq 0. \tag{3}$$

A probability measure $\pi(ds)$ over a compact set S governs an exogenous i.i.d. shock s_t that determines both government purchases and payoffs on the single security, positive random variables g, p with means $\bar{g}, \bar{p}$.

We let $s^t = (s_0, ..., s_t)$ denote a history of shocks. We often use x_t to denote a random variable x with a time t conditional distribution that is a function of history s^{t-1}. It is convenient to define $B_t \equiv q_t \mathrm{B}_t$ and $R_{t+1} \equiv p_{t+1}/q_t$ and to re-write the government's time t budget constraint (3) as

$$g_t + R_t B_{t-1} = \tau_t l_t + B_t.$$

A representative agent's time t budget constraint is

$$c_t + b_t = (1 - \tau_t)\, l_t + R_t b_{t-1}, \tag{4}$$

where b_t is his purchase of the single security. The period t market clearing condition for the security is

$$b_t = B_t. \tag{5}$$

We exogenously confine government debt to a compact set

$$B_t \in \left[\underline{B}, \overline{B}\right]. \tag{6}$$

The assumption of compactness of the feasible debt simplifies the analysis. We make the bounds sufficiently large that they do not affect the properties of the joint invariant distributions of government debt and the tax rate that we analyze below.

Definition 1. A *competitive equilibrium* given an initial government debt B_{-1} at $t = 0$ is a sequence $\{c_t, l_t, B_t, b_t, R_t, \tau_t\}_t$ such that (i) $\{c_t, l_t, b_t\}_t$ maximize (1) subject to the budget constraints (3); and (ii) constraints (2), (5), and (6) are satisfied. An *optimal competitive equilibrium* given B_{-1} is a competitive equilibrium that has the highest value of (1).

The single-security incomplete markets models of Barro (1979) and Aiyagari et al. (2002) assume that the security's payout is risk-free, a special case of our setup in which $p(s)$ is independent of s. Two purposes induce us to allow stochastic payoffs. First, standard real business cycle models driven by productivity and/or expenditure shocks fail to generate realistic holding period returns. Our way of modeling security markets allows us to remedy this shortcoming parsimoniously.[3] In Section 4 we document how post WWII real returns to U.S. government portfolios have fluctuated and use those findings to discipline payoffs. Second, as we show in Section 3.4, the optimal process for risk-free government debt when the representative consumer is risk-averse in consumption resembles the optimal behavior of state-contingent debt when the representative agent is risk-neutral. With risk-averse consumers and risk-free debt, a key object is an "effective return" on debt that takes into account a shadow cost of raising revenues; this influential return is stochastic even when the single security is risk-free. We show in Section 3.4 that the main insights of the quasilinear-stochastic-payoff formulation extend to economies with risk-averse consumers by considering "effective returns" on securities.

The representative consumer's first-order necessary conditions for an optimum imply that

$$1-\tau_t = l_t^{\gamma},\ \mathbb{E}_{t-1}R_t = \frac{1}{\beta}. \tag{7}$$

The security price q_t satisfies $q_t = \beta\bar{p}$, so the return on the security $R_t\left(s^t\right) = \frac{p(s_t)}{\beta\bar{p}}$. Substitute (7) into the consumer's budget constraint to obtain

$$c_t = l_t^{1+\gamma} + R_t B_{t-1} - B_t. \tag{8}$$

Use (8) to eliminate c_t from the feasibility condition (2) to obtain the following implementability constraints:

$$l_t - l_t^{1+\gamma} + B_t = R_t B_{t-1} + g_t. \tag{9}$$

Lemma 1. *$\{c_t, l_t, B_t, b_t, R_t, \tau_t\}_t$ is a competitive equilibrium given B_{-1} if and only if $\{l_t, B_{t-1}\}_t$ satisfies (6) and (9) for all $t \geq 0$.*

Lemma 1 allows us to compute an optimal competitive equilibrium allocation and a government debt process by solving

$$\max_{\{l_t, B_t\}_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(R_t B_{t-1} - B_t) + \frac{\gamma}{1+\gamma} l_t^{1+\gamma} \right], \tag{10}$$

where maximization is subject to constraints (6) and (9). The objective function in (10) is a version of (1) in which we have used (8) to eliminate c_t.

Appendix A.1 shows that it is optimal to set the tax rate to the left of the peak of the Laffer curve, which implies that the optimal tax rate τ_t and labor supply l_t are described by one-to-one mappings from total tax revenues $Z_t = \tau_t l_t$. Tax revenues are bounded from above by the level $\overline{Z}$ associated with a tax rate at the peak of the Laffer curve.[4] For a given level of tax revenues

[3] In the Appendix B we show that this set up is equivalent to risk-free bonds and discount factor shocks as in Albuquerque et al. (2016 forthcoming). We chose our stochastic payoff formulation because it naturally extends to multiple assets and portfolio theory that we study in Section 3.2.

[4] The expression for the maximum revenue is $\overline{Z} = \gamma\left(\frac{1}{1+\gamma}\right)^{1+1/\gamma}$.

Z, the corresponding tax rate $\tau(Z)$ and labor supply $l(Z)$ satisfy

$$\begin{aligned} Z &= \tau(Z)(1-\tau(Z))^{\frac{1}{\gamma}} \\ &= l(Z) - l(Z)^{1+\gamma}, \end{aligned} \tag{11}$$

which are well defined for all $Z \leq \overline{Z}$. Functions $l(\cdot), -\tau(\cdot)$ are decreasing. Let $\Psi(Z) \equiv \frac{1}{1+\gamma} l(Z)^{1+\gamma}$ be the utility cost of supplying labor required to raise tax revenues Z. Ψ is strictly decreasing, strictly concave, and differentiable on $(-\infty, \overline{Z}]$ and satisfies Inada conditions $\lim_{Z\to-\infty} \Psi'(Z) = 0$ and $\lim_{Z\to\overline{Z}} \Psi'(Z) = -\infty$.

An optimal value function $V(B\_)$ for problem (10) satisfies the Bellman equation

$$V(B\_) = \max_{Z(\cdot),B(\cdot)} \int \left[(R(s)B\_ - B(s)) + \gamma\Psi(Z(s)) + \beta V(B(s))\right]\pi(ds) \tag{12}$$

where maximization is subject to $Z(s) \leq \overline{Z}, B(s) \in [\underline{B}, \overline{B}]$, and

$$Z(s) + B(s) = R(s)B\_ + g(s) \quad \text{for all } s. \tag{13}$$

Strict concavity and differentiability of Ψ implies that V is also strictly concave and differentiable. Policy functions $\tilde{B}(s, B\_)$ and $\tilde{Z}(s, B\_)$ attain the right side of Bellman equation (12). Let $\tilde{\tau}(s, B\_)$ denote the associated optimal tax rate policy. Gross government expenditures $E(s, B\_)$, an important endogenous variable, are

$$E(s, B\_) = R(s)B\_ + g(s), \tag{14}$$

which equals government expenditures including interest and repayment of government debt. Aggregate shocks have effects on $E(s, B\_)$ that depend partly on government debt $B\_$.

We begin our analysis by stating a lemma that summarizes some key properties of optimal policy rules.

Lemma 2. *$\tilde{B}$, $\tilde{Z}$, and $\tilde{\tau}$ are increasing in E in the sense that $E(s'', B\_) > E(s', B\_)$ implies $\tilde{B}(s'', B\_) \geq \tilde{B}(s', B\_)$, $\tilde{Z}(s'', B\_) \geq \tilde{Z}(s', B\_)$, and $\tilde{\tau}(s'', B\_) \geq \tilde{\tau}(s', B\_)$ with strict inequalities if $\tilde{B}(s'', B\_), \tilde{B}(s', B\_) \in (\underline{B}, \overline{B})$.*

Let $\left\{\tilde{B}_t, \tilde{Z}_t\right\}_t$ be the optimum process generated by policy functions $\tilde{B}(s, B\_)$ and $\tilde{Z}(s, B\_)$. First-order conditions associated with the maximization problem (12) imply that if $\tilde{B}_t$ is interior, then the marginal social value of assets $V'(\tilde{B}_t)$ satisfies[5]

$$V'(\tilde{B}_t) = \beta\mathbb{E}_t R_{t+1} V'(\tilde{B}_{t+1}) = \mathbb{E}_t V'(\tilde{B}_{t+1}) + \beta \mathrm{cov}_t(R_{t+1}, V'(\tilde{B}_{t+1})). \tag{15}$$

Monotonicity of the policy functions asserted in Lemma 2 together with (15) allow us to prove:

Proposition 1. *The optimal process $\left\{\tilde{B}_t, \tilde{Z}_t\right\}_t$ has a unique invariant distribution.*

[5] Appendix A.1 provides an analysis of the situation in which $\tilde{B}_t$ is not required to be interior. Farhi (2010) obtained a generalized version of this equation in an economy with capital.

To enable us to characterize this invariant distribution, a key concept will be a level of debt

$$B^* \equiv \arg\min_B \operatorname{var}\big(RB + g\big) = -\frac{\operatorname{cov}(R, g)}{\operatorname{var}(R)}. \tag{16}$$

We assume that probability measure π is such that $B^* \in (\underline{B}, \overline{B})$ and that $\overline{B}$ is weakly below the natural debt limit. We call B^* the *risk-minimizing level of debt.* Let $Z^* \equiv \bar{g} + \frac{1-\beta}{\beta} B^*$ be the constant tax revenues that satisfy government's budget constraint on average if $B_t = B^*$ for all t.

2.1 Perfectly correlated shocks: the exact characterization

We first consider a special case in which p and g are perfectly correlated that illustrates key economic forces that determine the long-run behavior of debt and taxes more generally.

Proposition 2. *Suppose that p and g are perfectly correlated. Then $\tilde{B}_t \to B^*$, $\tilde{Z}_t \to Z^*$ a.s.*

Proof. If p and g are perfectly correlated then $\operatorname{cov}(E(\cdot, B), R(\cdot)) \geq 0$ if $B \geq B^*$, $\operatorname{cov}(E(\cdot, B), R(\cdot)) \leq 0$ if $B \leq B^*$, and $E(s, B)$ is independent of s if and only if $B = B^*$. The monotonicity of policy functions established in Lemma 2 and concavity of V imply that $\operatorname{cov}_t(R_{t+1}, V'(\tilde{B}_{t+1})) \leq 0$ if $\tilde{B}_t \geq B^*$ and $\operatorname{cov}_t(R_{t+1}, V'(\tilde{B}_{t+1})) \geq 0$ if $\tilde{B}_t \leq B^*$. Therefore the martingale convergence theorem and equation (15) imply that $V'(\tilde{B}_t)$ converges almost surely. By strict concavity, it can converge only to a level of debt B for which $E(s, B)$ is independent of s, which is only possible if $\tilde{B}_t \to B^*$ a.s. Since $\mathbb{E}R_t = \beta^{-1}$, equation (13) establishes that $\tilde{Z}_t \to Z^*$ a.s. □

An insight of Proposition 2 is that the conditional covariance in equation (15) induces a drift in the stochastic process $\tilde{B}_t$ towards the risk-minimizing level of debt B^*. Here is some intuition. Fluctuations in tax rates, and therefore tax revenues, have welfare costs for reasons explained by Barro (1979). For this reason, on the margin each period the planner wants to move closer to a risk-minimizing level of debt that reduces his need to change the tax rate in response to shocks to government purchases. When p and g are perfectly correlated, fluctuations in returns on government debt $R(s)B^*$ perfectly offset fluctuations in government expenditures $g(s)$, thereby providing a perfect hedge. In this situation, the tax rate τ_t is constant in the long run.

2.2 General case: approximations

When p and g are imperfectly correlated, perfect hedging is impossible. To study this case, we develop a class of second order approximations that do a good job of approximating the joint invariant distribution of government debt and tax revenues. Under particular conditions, our approximating policies are *linear* in shocks, a property that facilitates asymptotic analysis.

We start with the observation that random variables g and p can be expressed as

$$g(s) = \bar{g} + \sigma\epsilon_g(s), \quad p(s) = \overline{p} + \sigma\epsilon_p(s),$$

where ϵ_g and ϵ_p have mean zero but can be arbitrarily correlated with each other. We will study the properties of a Ramsey plan when shocks are *small,* i.e. as σ approaches to zero. Let $\tilde{B}\left(s,B\_;\sigma\right)$ and $\tilde{Z}\left(s,B\_;\sigma\right)$ be policy functions for a given σ. Optimality conditions for problem (12) should hold for all realizations of $p(s), g(s)$ and for all values of σ. Therefore first, second, and higher order derivatives of those optimality conditions with respect to $\epsilon_g, \epsilon_p, \sigma$, assuming they exist, must all be equal to zero.[6] That insight allows us to calculate the Taylor expansion of policy rules around a current level of debt since $\tilde{B}(s,B\_;0)=B\_$. In Appendix A.1, we show that[7]

$$\begin{aligned}\tilde{B}\left(s,B\_\right) &= B\_ + \beta\left[g(s)-\bar{g}\right] + \left[\beta R(s)-1\right]B\_ - \beta^2\mathrm{var}(R)B\_ - \beta^2\mathrm{cov}(R,g)\\ &+ \mathcal{O}(\sigma^3,(1-\beta)\sigma^2). \qquad (17)\end{aligned}$$

The second order expansion is linear in g and R up to terms that appear in $\mathcal{O}(\cdot)$. Since standard macroeconomic calibrations set the discount factor close to 1, we drop the $\mathcal{O}(\cdot)$ terms and proceed to study an optimal debt and tax policy implied by that approximation.[8] The linearity of the policy rules allows us to obtain a simple and transparent characterization. We show later in this section and in Section 4 that this procedure provides good approximations to other more accurate approximations computed using global numerical algorithms and has the virtue of shedding light on economic principles underlying optimal debt and tax policies.

We focus on three moments: means, variances, and speeds of mean reversion to the invariant distribution of debt and taxes. We obtain these by re-grouping terms in equation (17) and integrating with respect to the ergodic measure. For example, by taking unconditional expectations on both sides of (17), we deduce that the unconditional mean and variance of debt can be estimated up to $\mathcal{O}(\sigma,(1-\beta))$ terms.[9]

Proposition 3. *The invariant distribution of* $\left\{\tilde{B}_t,\tilde{Z}_t\right\}_t$ *satisfies*

- **Means**
$$\mathbb{E}\left(\tilde{B}_t\right) = B^* + \mathcal{O}(\sigma,(1-\beta)), \quad \mathbb{E}\left(\tilde{Z}_t\right) = Z^* + \mathcal{O}(\sigma,(1-\beta));$$

- **Speeds of reversion to means**
$$\frac{\mathbb{E}_t\left(\tilde{B}_{t+1}-B^*\right)}{\tilde{B}_t-B^*} = \frac{\mathbb{E}_0\left(\tilde{Z}_{t+1}-Z^*\right)}{\mathbb{E}_0\left(\tilde{Z}_t-Z^*\right)} = \frac{1}{1+\beta^2\mathrm{var}\left(R\right)} + \mathcal{O}(\sigma^3,(1-\beta)\sigma^2);$$

- **Variances**
$$\mathrm{var}\left(\tilde{B}_t\right) = \frac{\mathrm{var}\left(RB^*+g\right)}{\mathrm{var}\left(R\right)} + \mathcal{O}(\sigma,(1-\beta)), \quad \mathrm{var}\left(\tilde{Z}_t\right) = \left(\frac{1-\beta}{\beta}\right)^2\mathrm{var}\left(\tilde{B}_t\right) + \mathcal{O}(\sigma^3,(1-\beta)\sigma^2).$$

[6] This approach is originally developed by Fleming (1971) and was applied in economics by Schmitt-Grohe and Uribe (2004). Like them, we assume that policies are twice differentiable.

[7] Here $\mathcal{O}(\sigma^3,(1-\beta)\sigma^2)$ denote all terms that appear as $\mathcal{O}(\sigma^3)$ or $(1-\beta)\,\mathcal{O}(\sigma^2)$.

[8] This approximation should work well so long as average interest rates are of similar or smaller order of magnitude than the standard deviation of shocks that affect government's budget constraint. This condition holds in our calibration to the post WWII U.S. data in Section 4.

[9] Observe that while $\mathrm{var}(R), \mathrm{var}(g), \mathrm{cov}(R,g)$ are all of order $\mathcal{O}(\sigma^2)$, functions $\frac{\mathrm{cov}(R,g)}{\mathrm{var}(R)}$ and $\frac{\mathrm{var}\left(RB^*+g\right)}{\mathrm{var}(R)}$ are of order $\mathcal{O}(1)$.

The first part of Proposition 3 shows that the risk-minimizing debt B^* is the mean of the invariant distribution, and the mean level of tax revenues is Z^*. To understand the finding that the mean of the invariant distribution of $\tilde{B}_t$ is B^*, it is useful to connect the martingale (15) to the static variance minimization problem (16). By strict concavity of the value function V, there is a one-to-one relationship between debt B_t and its marginal value to the planner, $V'(B_t)$. Inspection of the martingale equation (15) shows that the covariance term $\text{cov}_t(V'(B_{t+1}), R_{t+1})$ is important in determining the drift of the dynamics of debt in the long run. For a given B_t, the debt next period B_{t+1} depends only on E_{t+1} and consequently

$$\text{cov}_t(V'(B_{t+1}), R_{t+1}) \propto \text{cov}_t(E_{t+1}, R_{t+1}) + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2). \tag{18}$$

It is possible to verify that $\text{cov}_t(E_{t+1}, R_{t+1}) = \frac{1}{2}\frac{\partial \text{var}(R_{t+1}B_t+g_{t+1})}{\partial B_t}$. Thus, ignoring $\mathcal{O}(\sigma^3, (1-\beta)\sigma^2)$ terms, the covariance term in the martingale equation (15) is proportional to the slope of the variance of E_t with respect to government debt B_t. Since B^* minimizes variation in $E(s, B\_)$, the slope is zero at B^*. The change in signs of the slope implies that, to second order, $V'(B_t)$ is a submartingale when $B_t > B^*$ and supermartingale when $B_t < B^*$. Then arguments used in the proof of Proposition 2 explain why $\tilde{B}_t$ drifts towards B^*.

Proposition 3 also shows that the speed of mean reversion is determined by the variance of returns: a lower variance of returns decreases the speed of the reversion. When $B_t \neq B^*$ the fluctuations in the rate of return put additional risk into $E(s, B_t)$ that is increasing in the volatility of R and the magnitude of B_t. A more volatile R implies that it is optimal increase the speeds at which the government should repay debt when $B_t > B^*$ and should accumulate debt when $B_t < B^*$. Dynamics of debt and taxes both approximate random walks when the security is nearly a risk-free bond, confirming an insight of Barro (1979).

The last part of Proposition 3 characterizes second moments of the invariant distribution of debt and tax tax rates. It shows several insights. First, the dispersion of the invariant distribution of government debt is increasing in un-hedgable risk as measured by $\frac{\text{var}(RB^*+g)}{\text{var}(R)}$. Note that this terms does not depend on σ (i.e. it is $O(1)$) and is zero only when g and p are perfectly correlated. So long as g and p are imperfectly correlated, the variance of the invariant distribution of debt does not vanish even when σ becomes small. This outcome reflects two offsetting forces: smaller shocks imply that debt reacts less to arrival of a shock, but also that it takes longer for debt to revert to its mean. The variance of tax revenues $\tilde{Z}_t$ is determined by two considerations. Tax revenues must respond enough to changes in the level of government debt to satisfy the budget constraint, which is captured by the term $\left(\frac{1-\beta}{\beta}\right)^2 \text{var}\left(\tilde{B}_t\right)$. Tax revenues also change in response directly to an expenditure shock. Since the planner wants to smooth tax rates over time, only a fraction $1-\beta$ of an innovation to government expenditures is financed by contemporaneous changes in tax revenues. Therefore, the variance induced by a contemporaneous response to aggregate shocks is of order $\mathcal{O}(\sigma^3, (1-\beta)\sigma^2)$.

Figure I illustrates the accuracy of the quadratic approximation. As a baseline we set $\beta = 0.98$ and $\gamma = 2$ and choose the joint stochastic processes for (g, p) to match the standard deviation of government expenditures, returns on government's debt portfolio and the correlation between these returns and government net-of-interest expenditures. The upper bound $\overline{B}$ is chosen to be equal to the natural debt limit that we can compute explicitly for the quasilinear setup and the lower bound is set so that the debt-to-output ratio is approximately -300%. For all the exercises we report below, we verify that $B^* \in [\underline{B}, \overline{B}]$.[10] The g, p processes are modeled as

$$\begin{aligned} g(s) &= \bar{g} + \sigma\epsilon_g(s), \\ p(s) &= 1 + \chi\sigma\epsilon_g(s) + \sigma\epsilon_{\hat{p}}(s), \end{aligned}$$

where $\sigma = 1$ and the shocks $\epsilon_{\hat{p}}, \epsilon_g$ are finite state approximations to standardized normal random variables.[11] The moments we target in addition to the parameter values that achieve those targets are reported in Table I.

Given these primitives, we compute the optimal policies from first order conditions for (12), and iterating on the planner's Euler equation using cubic splines as basis functions for approximating policies.[12] We then compare the outcomes of our global solution to the quadratic approximations. We plot the invariant distribution of debt and policy rules obtained from the global solution method (dashed lines) and the quadratic approximations (solid lines) in Figure I. For parsimony, we plot policies $\tilde{B}(s, B\_) - B\_$ for two values of s that correspond to the smallest and the largest pairs of $(g(s), p(s))$. The top panel of Figure I reveals that the ergodic distribution of debt obtained from the quadratic approximations of policies closely resemble policies obtained using the global numerical method. The top right panel reveals see that, as current debt approaches the natural debt limit, the quadratic approximation to the optimal government debt policy does not capture the proper curvature with respect to B. However, at the parameters that we used to compute plots in the top panel, the ergodic distribution puts almost no mass on regions where the quadratic approximations and global approximations differ.

Proposition 3 states that our approximation errors increase with $1 - \beta$ and σ. To check how quickly these approximation errors become significant, we reduce β to 0.90 and increase σ to 4 in the second and third rows of Figure I respectively. For most of the state space, we find that the quadratic approximation continues to do well. As a consequence of the fact that the our quadratic approximations assume interiority, the policies reported in the right panel display approximation errors only when debt approaches the debt limits. When $1 - \beta$ or σ is high, the quadratic approximations imply slightly higher debt than does the solution computed with numerical value function iteration. Almost all of these differences emerge because the quadratic expansion puts positive probability on the region where debt is higher than $\bar{B}$.

3 Extensions

Forces isolated within the Section 2 economy prevail under alternative assumptions about motives for taxation, persistence in g and p and also fluctuations in productivity, rich sets of securities, and preferences that express aversion to consumption risk. We discuss these now.

[10] In Section 4 we do a comprehensive calibration where we match several moments of returns, output and debt to U.S. post war data for a richer model that allows for persistence, risk aversion, productivity shocks. Here we use a subset of those moments to get a reasonable baseline which when we modify in several directions to test the accuracy of our approximations. The details of the sample and data series used to construct these moments are in Section 4.

[11] The finite state approximation ensures that $g(s) > 0$ and $p(s) > 0$ for all s.

[12] Since the problem is concave such a fixed point corresponds to the optimal policies.

3.1 Transfers and redistribution

Optimal debt management in our Section 2 model differs significantly from that in other incomplete markets models studies by Aiyagari et al. (2002) and Farhi (2010). A key difference is that we prohibit lump-sum taxes or transfers, while Aiyagari et al. (2002) and Farhi (2010) allow positive but not negative lump-sum transfers. In our model, the invariant joint distribution of debt and taxes is unique. In the long run, debt and tax rates minimize fluctuations in gross government expenditures including debt service requirements, $E(s, B)$. By way of contrast, optimal plans in Aiyagari et al. (2002) and Farhi (2010) have a continuum of invariant distributions of debt levels. In all of them, tax rates are zero and debt levels are negative and big enough in absolute value to finance all net-of-interest government expenditures from earnings on the government's portfolio, and fluctuations in transfers fully absorb shocks to net-of-interest government expenditures. Here we extend our analysis to an economy with lump sum transfers by explicitly modeling the utility enjoyed by recipients of transfers. We show that our Section 2 results carry over essentially unchanged so long as the utility function of recipients of transfers is strictly concave. In that case, uncertainty about transfers is costly, prompting the government to use government debt to minimize fluctuations in both tax rates and transfers. We then discuss what drives the long-run tax rate to zero in Aiyagari et al. (2002) and explain how to reconcile their results with ours.

A standard justification for ruling out lump-sum taxes in representative agent models is implicitly to appeal to the presence of unmodeled "poor" agents who cannot afford to pay a lump-sum tax. In this section, we study optimal anonymous transfers in an economy with such poor agents. We extend the Section 2 economy to have just enough heterogeneity across agents to make the analysis meaningful. In particular, we assume that in addition to a measure 1 of agents of type 1 with quasilinear preferences $U(c, l) = c - \frac{l^{1+\gamma}}{1+\gamma}$, there is a measure $n > 0$ of type 2 agents who cannot work or trade securities and who enjoy utility

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U\left(c_{2,t}\right),$$

where $c_{2,t}$ is consumption of a type 2 agent in period t; U is strictly concave and differentiable on $\mathbb{R}_+$ and satisfies the Inada condition $\lim_{c \to 0} U'(c) = \infty$.

The government and type 1 agent trade the Section 2 security. The government imposes a linear tax rate τ_t on labor income and awards lump-sum transfers T_t that do not depend on the type of agent. Negative transfers are not feasible because a type 2 agent has no income other than transfers. Each agent receives a per-capita transfer $\frac{T_t}{1+n}$. Since agent 2 lives hand to mouth, his budget constraint is

$$c_{2,t} = \frac{T_t}{1+n}.$$

The planner ranks allocations according to

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\left(c_t - \frac{1}{1+\gamma} l_t^{1+\gamma} \right) + \omega U\left(c_{2,t}\right) \right],$$

for some $\omega > 0$.

The time t government budget constraint is now

$$g_t + T_t + R_t B_{t-1} = \tau_t l_t + B_t.$$

With only minimal modifications, the budget constraint of a type 1 agent, definition 1 of a competitive equilibrium, and the Section 2 recursive formulation of the optimal policy problem all extend to this environment. The planner's optimal value function satisfies the Bellman equation

$$\begin{aligned} V\left(B\_\right) = \max_{Z(\cdot),B(\cdot),T(\cdot)} \int \Bigg[\left(R\left(s\right) B\_ - B\left(s\right) + \frac{T\left(s\right)}{1+n} \right) + \gamma \Psi\left(Z\left(s\right)\right) \\ + \omega U\left(\frac{T\left(s\right)}{1+n}\right) + \beta V\left(B\left(s\right)\right) \Bigg] \pi\left(ds\right) \end{aligned} \quad (19)$$

subject to $Z(s) \leq \overline{Z}, B\left(s\right) \in [\underline{B}, \overline{B}]$, and

$$Z\left(s\right) - T\left(s\right) + B\left(s\right) = R\left(s\right) B\_ + g\left(s\right) \text{ for all } s. \quad (20)$$

Denoting by $\left\{\tilde{B}_t, \tilde{Z}_t, \tilde{T}_t\right\}_t$ the outcomes associated with policies that attain $V\left(B\_\right)$ and following the same steps as in the proofs of Section 2, we obtain

Proposition 4. *The invariant distribution of* $\left\{\tilde{B}_t, \tilde{Z}_t, \tilde{T}_t\right\}_t$ *is unique. The invariant distribution of* $\tilde{B}_t$ *satisfies properties stated in Proposition 3. The invariant distribution of* $\tilde{Z}_t - \tilde{T}_t$ *has the same properties as the invariant distribution of* $\tilde{Z}_t$ *in Proposition 3. Let* $F(\tilde{T};\omega)$ *be the cumulative distribution function of the ergodic distribution of* $\tilde{T}_t$. *If* $\omega > \omega'$ *then* $F(\tilde{T};\omega)$ *first order stochastically dominates* $F(\tilde{T};\omega')$.

Insights from Section 2 about optimal debt management carry over to this heterogeneous economy. Fluctuations in the tax rate and (non-agent-specific) lump-sum transfers now are both costly, so an optimal policy smooths both. Adjusting the tax rate in response to government expenditure shocks is costly because the dead-weight loss of taxation is convex in tax rates, as stressed by Barro (1979). Adjusting transfers is also costly because that induces fluctuations in inequality.

In Aiyagari et al. (2002) and Farhi (2010) the government eventually sets tax rates to zero and thereafter adjusts transfers one-to-one with government expenditures. They do not model heterogeneity explicitly but appeal to it only implicitly when they impose $T_t \geq 0$. That restriction puts a kink in the cost of using transfers: for sufficiently large government assets, the marginal cost of an increase in transfers is zero, while the marginal cost of a decrease in transfers is infinite at $T_t = 0$. A high marginal cost of negative transfers creates an incentive for the governments in Aiyagari et al. (2002) and Farhi (2010) to accumulate enough assets to make the constraint $T_t \geq 0$ eventually become slack. Since fluctuations in positive transfers are costless, in the long-run the government uses those transfers to offset all fluctuations in expenditures g_t.

By way of contrast, in our economy, the welfare cost of using transfers is endogenous and smooth, so that marginal costs from increasing and decreasing transfers around an optimal level $\tilde{T}_t$ are the same, $\frac{\omega}{1+n} U'\left(\tilde{T}_t\right)$; welfare costs of departing from the optimal inequality level are strictly convex. This difference accounts for the very different long run dynamics than those discovered by Aiyagari et al. (2002).[13]

The restriction that transfers, T_t are common across all types of agents is not essential for Proposition 4. Consider a slightly modified taxation scheme where the government uses a linear tax rate, meaning one with a zero intercept, for the productive type of agent and a lump sum transfer for the unproductive types. The budget constraint of type 2 is $c_{2,t} = \frac{T_t}{n}$ and the Bellman equation (19) is altered to

$$V\left(B\_\right) = \max_{Z(\cdot),B(\cdot),T(\cdot)} \int \left[\left(R\left(s\right)B\_ - B\left(s\right)\right) + \gamma\Psi\left(Z\left(s\right)\right) + \omega U\left(\frac{T\left(s\right)}{n}\right) + \beta V\left(B\left(s\right)\right)\right]\pi\left(ds\right).$$

We show in Appendix A.2 that Proposition 4 continues to hold. While the assumption that only unproductive agents receive transfers changes the average level of optimal tax revenues and the tax rate, it leaves unaffected the moments of the Ramsey policy characterized in Proposition 4.

3.2 More general asset structure

In this section, we study optimal management of a government's portfolio of securities by modifying the baseline Section 2 setup to allow $K \geq 1$ securities. Let $p^k(s)$ be the payoff of security k in state s. Each security is available in fixed net supply Q^k, which can be nonzero. Our setup thus allows for trade in financial assets like government debt and claims to Lucas trees. When Lucas trees are available, the feasibility constraint reads

$$c_t + g_t = l_t + \sum_{k=1}^{K} p_t^k Q^k.$$

To simplify, we assume that available securities consist of one period lived securities and infinitely lived consols, and that s is an i.i.d. process.[14] Let B_t^k be the government's holdings of security k at the end of period t, q_t^k its market price, and ι^k an indicator variable that is equal to 1 if security k is a consol. The government's time t budget constraint is

$$g_t + \sum_{k=1}^{K}\left(p_t^k + \iota^k q_t^k\right)\mathrm{B}_{t-1}^k = \tau_t l_t + \sum_{k=1}^{K} q_t^k \mathrm{B}_t^k.$$

Let $R_t^k = \frac{p_t^k + \iota^k q_t^k}{q_{t-1}^k}$ be the gross return on security k and let $B_t^k = q_t^k \mathrm{B}_t^k$ be the market value of holdings of security k so that we can write this budget constraint as

$$g_t + \sum_{k=1}^{K} R_t^k B_{t-1}^k = \tau_t l_t + \sum_{k=1}^{K} B_t^k.$$

Let $B_t \equiv \sum_{k=1}^{K} B_t^k$ be the market value of the government's portfolio, which we restrict to be in a compact set $\left[\underline{B}, \overline{B}\right]$. We assume that these bounds are sufficiently large so that the risk-

[13] Bhandari et al. (2015b) show that this insight carries over to richer economies with much more heterogeneity and in which no agent is excluded from the financial markets.

[14] The extension of our results to arbitrary finite period securities is straightforward but requires additional notation. Extensions to richer shock processes will follow along the lines of Section 3.3.

minimizing portfolio B^* to be defined below is feasible. Without loss of generality, we assume that no security is redundant in the sense that the vectors $\left\{R^k\right\}_{k=1}^K$ are linearly independent. We use $\mathbf{R}(s)$ to denote returns $\left(R^1(s),...,R^K(s)\right)$, $\mathbf{B}$ and $\mathbf{1}$ to denote a K dimensional (column) vector of security holdings and of ones, respectively. Let $\mathbb{C}[\mathbf{R},\mathbf{R}]$ and $\mathbb{C}[\mathbf{R},g]$ be a matrix of the covariances of returns and a vector of covariances of returns with government purchases g, respectively. When the matrix $\mathbb{C}[\mathbf{R},\mathbf{R}]$ is nonsingular, we define

$$B^* \equiv -\mathbf{1}^\intercal \mathbb{C}[\mathbf{R},\mathbf{R}]^{-1}\mathbb{C}[\mathbf{R},g], \tag{21}$$

which, as we show below, is the risk-minimizing level of government debt that generalizes equation (16) to the case of multiple assets. Whenever B^* is well defined, we also define $Z^* \equiv \frac{1-\beta}{\beta}B^* + \bar{g}$.

Temporarily suppose that government portfolio weights are fixed, meaning that there exist constants $\psi_1,...,\psi_K$ such that $\frac{B_t^k}{\sum_k B_t^k} = \psi_k$ for all t. Define $R(s) \equiv \sum_{k=1}^K \psi_k R^k(s)$. Then the optimal policy problem is equivalent to the one in Section 2. Thus, if the government arbitrarily fixes its portfolio weights then, subject to that arbitrary choice, all Section 2 insights about optimal debt management and fiscal policy still apply.

Now suppose that the Ramsey planner optimally chooses government portfolio weights each period. The Ramsey problem in this case can be written recursively. The end of period market value of the government's portfolio the only state variable in the planner's value function:

$$V(B\_) = \max_{Z(\cdot),B(\cdot),\mathbf{B}} \int \left[\left(\mathbf{R}(s)^T\mathbf{B} - B(s)\right) + \gamma\Psi(Z(s)) + \beta V(B(s))\right]\pi(ds) \tag{22}$$

where maximization is subject to $Z(s) \leq \overline{Z}, B(s) \in [\underline{B},\overline{B}]$, $\mathbf{1}^\intercal\mathbf{B} = B\_$, and

$$B(s) + Z(s) = \mathbf{R}(s)^T\mathbf{B} + g(s) \quad \text{for all } s. \tag{23}$$

We first establish that:

Lemma 3. *Problem (22) has a unique solution. If $\mathbb{C}[\mathbf{R},\mathbf{R}]$ is nonsingular or if g is not in the span of $\mathbf{R}$, then the invariant distribution generated by policies that attain $V(B\_)$ is unique; otherwise optimal policies satisfy $\tilde{B}(s,B\_) = B\_$ for all s and $Z(s,B\_)$ is independent of s.*

As in the Section 2 baseline model, we scale the volatility of all shocks shocks by σ and take a second order Taylor expansion of the policies $\tilde{Z}(s,B\_;\sigma)$, $\tilde{B}(s,B\_;\sigma)$, and $\tilde{\mathbf{B}}(B\_;\sigma)$ around $\sigma = 0$. At $\sigma = 0$ the portfolio problem is indeterminate, but the next lemma shows that there is a unique limiting portfolio as $\sigma \to 0$ that solves a variance minimization problem.

Lemma 4. $\lim_{\sigma\to 0}\tilde{\mathbf{B}}(B\_;\sigma) = \mathbf{B}^*(B\_)$ *where* $\mathbf{B}^*(B\_)$ *solves*

$$\min_{\mathbf{B}} \operatorname{var}\left(\sum_k B^k R^k + g\right) \quad \textit{subject to} \quad \mathbf{1}^\intercal\mathbf{B} = B\_. \tag{24}$$

We can relate B^* in equation (21) to $\mathbf{B}^*(\cdot)$ defined in Lemma 4. Suppose that the matrix $\mathbb{C}[\mathbf{R},\mathbf{R}]$ is non-singular. Then the constraint $\mathbf{1}^\intercal\mathbf{B} = B\_$ in problem (24) binds, making the variance depend on the level of debt $B\_$. The variance is minimized at $B\_ = B^*$, making B^* the risk-minimizing debt level satisfying $B^* = \mathbf{1}^\intercal \mathbf{B}^* (B^*)$.[15]

As in Section 2, we show that the B^* is the long-run mean of the second-order approximation to the optimal policy for the market value of the government debt portfolio. When $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is nonsingular, the Taylor expansion around $\mathbf{B}^* (B\_)$ yields

$$\begin{aligned}\tilde{B}(s, B\_) &= B\_ + \beta[g(s) - \bar{g}] + [\beta \mathbf{R}(s) - \mathbf{1}]^T \mathbf{B}^* (B\_) \\ &\quad - \frac{\beta^2 B\_}{\mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}} - \frac{\beta^2 \mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbb{C}[\mathbf{R}, g]}{\mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}} + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2).\end{aligned}$$

When $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is singular,

$$\tilde{B}(s, B\_) = B\_ + \beta[g(s) - \bar{g}] + \beta[\mathbf{R}(s) - \mathbf{1}]^T \mathbf{B}^* (B\_) + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2). \tag{25}$$

Proposition 5. *Suppose that $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is non-singular. The invariant distribution of $\left\{\tilde{B}_t, \tilde{Z}_t\right\}_t$ has*

- ***Means***

$$\mathbb{E}\left(\tilde{B}_t\right) = B^* + \mathcal{O}(\sigma, (1-\beta)), \quad \mathbb{E}\left(\tilde{Z}_t\right) = Z^* + \mathcal{O}(\sigma, (1-\beta)).$$

- ***Speeds of mean reversions***

$$\frac{\mathbb{E}_t\left(\tilde{B}_{t+1} - B^*\right)}{\tilde{B}_t - B^*} = \frac{\mathbb{E}_0\left(\tilde{Z}_{t+1} - Z^*\right)}{\mathbb{E}_0\left(\tilde{Z}_t - Z^*\right)} = \frac{\beta^{-2} \mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}}{1 + \beta^{-2} \mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}} + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2),$$

- ***Variances***

$$\begin{aligned}\operatorname{var}\left(\tilde{B}_t\right) &= \left(\mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}\right) \operatorname{var}\left(-\mathbb{C}[\mathbf{R}, g]^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{R} + g\right) + \mathcal{O}(\sigma, (1-\beta)), \\ \operatorname{var}\left(\tilde{Z}_t\right) &= \left(\frac{1-\beta}{\beta}\right)^2 \operatorname{var}\left(\tilde{B}_t\right) + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2).\end{aligned}$$

Government holdings of individual securities satisfy

$$\tilde{\mathbf{B}}_t = -\mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbb{C}[\mathbf{R}, g] + \frac{\mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}}{\mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbf{1}} \left(\tilde{B}_t + \mathbf{1}^\intercal \mathbb{C}[\mathbf{R}, \mathbf{R}]^{-1} \mathbb{C}[\mathbf{R}, g]\right) + \mathcal{O}(\sigma, (1-\beta)). \tag{26}$$

Some examples illustrate these findings.

Example 1. Suppose that there are two securities with $\operatorname{var}\left(R^k\right) > 0$ for $k = 1, 2$ and that the return on security 1 is perfectly correlated with g while the return on security 2 is orthogonal to the return on security 1. Then Proposition 5 implies that the ergodic mean of the value of government's debt portfolio is $B^* = -\frac{\operatorname{cov}\left(R^1, g\right)}{\operatorname{var}\left(R^1\right)}$, that the speed of convergence to B^* is $\left(1 + \beta^2 \frac{\operatorname{var}\left(R^2\right)}{\operatorname{var}\left(R^1\right) + \operatorname{var}\left(R^2\right)} \operatorname{var}\left(R^1\right)\right)^{-1}$, and that its ergodic variance is zero. From formula (26),

[15] If $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is singular, then the constraint $\mathbf{1}^\intercal \mathbf{B} = B\_$ does not bind and any debt level is risk-minimizing.

the optimal portfolio along transition paths satisfies

$$\tilde{B}^1(\tilde{B}_t) = \frac{\mathrm{var}\left(R^2\right)}{\mathrm{var}\left(R^1\right) + \mathrm{var}\left(R^2\right)}\tilde{B}_t + \frac{\mathrm{var}\left(R^1\right)}{\mathrm{var}\left(R^1\right) + \mathrm{var}\left(R^2\right)}B^*$$

$$\tilde{B}^2(\tilde{B}_t) = \frac{\mathrm{var}\left(R^1\right)}{\mathrm{var}\left(R^1\right) + \mathrm{var}\left(R^2\right)}\tilde{B}_t - \frac{\mathrm{var}\left(R^1\right)}{\mathrm{var}\left(R^1\right) + \mathrm{var}\left(R^2\right)}B^*,$$

with $\tilde{B}^2(\tilde{B}_t) \to 0$ a.s.

Complete hedging can be achieved with the government holding security 1 only, just as in Proposition 2, so that holding any security 2 is suboptimal asymptotically. If the market value of the initial government debt does not equal B^*, it is optimal to invest in security 2 along the transition path because doing this reduces risk for the government until the steady state is reached. As a result, noting that $\frac{\mathrm{var}\left(R^2\right)}{\mathrm{var}\left(R^1\right)+\mathrm{var}\left(R^2\right)} < 1$, the speed of convergence to the long-run portfolio is slower than when only security 1 can be traded.

Example 2. Consider a setting with two securities which payoffs are perfectly correlated with g and $0 \leq \mathrm{var}(R^1) < \mathrm{var}(R^2)$.[16] There exist unique constants ψ_1, ψ_2, ξ_1, and ξ_2 such that

$$\psi_1 R^1(s) + \psi_2 R^2(s) = g(s)$$

and

$$\xi_1 R^1(s) + \xi_2 R^2(s) = \frac{1}{\beta}.$$

Note that $\psi_1 + \psi_2 = \beta\bar{g}$ and $\xi_1 + \xi_2 = 1$. Now the covariance matrix $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is singular. The risk-minimizing portfolio satisfies $B^{*,k}(B\_) = (B\_ + \beta\bar{g})\,\xi_k - \psi_k$. Holding it allows the government to attain complete markets allocations for any $B\_$; the value of government debt equal its initial value for all $t \geq 0$.

It is instructive to study how an optimal portfolio changes as R^2 approaches R^1. For simplicity, suppose that $R^1(s) = \frac{1}{\beta}$ and $R^2(s) = \frac{1}{\beta} - \varepsilon(g(s) - \bar{g})$. For a given $B\_$, optimal asset positions are $B^{*,2} = \frac{1}{\varepsilon}$ and $B^{*,1} = B\_ - B^{*,2}$, both of which become arbitrarily large as $\varepsilon \to 0$. This outcome explains why Buera and Nicolini (2004) and Farhi (2010) found that the government should take extremely large asset positions to hedge its risk. Those papers allowed a planner to trade a risk-free one period security plus other securities (long bonds in Buera and Nicolini (2004), capital in Farhi (2010)). The returns on those securities had low volatilities and high correlations with government expenditures. Consistent with our example, those authors found that an optimal portfolio has huge positions in these securities.

Example 3. Suppose that $\mathrm{cov}\left(R^k, R^l\right) = 0$ for all $k \neq l$. Now $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is a diagonal matrix and $\frac{\partial \tilde{B}^k(B\_)}{\partial B\_} \propto \frac{1}{\mathrm{var}\left(R^k\right)}$ from (26). As the value of the outstanding government debt increases, its optimal composition shifts towards securities that have lower variances of returns. In the limit, as the variance of returns of one security approaches zero, all of the adjustments to changes in $B\_$ use that security. We can use this example to construct a simple model of an optimal

[16] Note that risk-free returns are in the closure of the set of returns that are perfectly correlated with g. We follow a convention of calling a risk-free security to be perfectly correlated with g.

maturity structure of government debt. Suppose that the government can issue a one-period risk-free bond and a consol with a stochastic coupon. Then the optimal issue of the consol is $\tilde{B}_t^2 = -\frac{\text{cov}(R^2,g)}{\text{var}(R^2)}$, which is independent of $\tilde{B}_t$, while the optimal issue of the riskless security is $\tilde{B}_t^1 = \tilde{B}_t - \tilde{B}_t^2$. Hence, the optimal effective maturity $\frac{\tilde{B}_t^2}{\tilde{B}_t^1}$ of government debt is decreasing in the value of outstanding debt $\tilde{B}_t$.

3.3 More general shock processes

In this section, we modify the Section 2 baseline model setup to include richer shock processes. In addition to expenditure and payoff shocks, we introduce fluctuations in productivity θ and allow g, p, θ to be correlated across time and with each other.

We follow the set up of Section 2 but assume that state $s = (p, g, \theta)$ follows a first order Markov process. The conditional probability density of s_t is described by a Markov kernel $\pi(\cdot|s\_)$, where $s\_$ is the realization of the shock in period $t-1$. We assume that π has a unique invariant measure λ. The feasibility constraint now takes the form

$$c_t + g_t = \theta_t l_t \tag{27}$$

and the return in state s is $R(s, s\_) = \frac{p(s)}{\beta \int p(s')\pi(ds'|s\_)}$.

Let $\Theta \equiv \theta^{\frac{1+\gamma}{\gamma}}$ and $Z \equiv \tau(1-\tau)^{\frac{1}{\gamma}}$. As in Section 2, there is a one to one correspondence between Z and τ for $Z \leq \bar{Z}$. The tax revenues with productivity shocks are equal to ΘZ. Let $\bar{g}$ and $\bar{\Theta}$ denote ergodic means of g and Θ. Let $\Omega(Z, s) \equiv \frac{l^{1+\gamma}(Z,s)}{1+\gamma}$, where the function $l^{1+\gamma}(Z, s)$ is now defined as

$$\Theta(s)Z = \Theta(s)^{\frac{\gamma}{1+\gamma}} l(Z, s) - l^{1+\gamma}(Z, s).$$

Following Section 2 arguments, the Ramsey planner's value function satisfies the Bellman equation

$$V(B\_, s\_) = \max_{Z(\cdot), B(\cdot)} \int \left[(R(s, s\_)B\_ - B(s)) + \gamma\Omega(Z(s), s) + \beta V(B(s), s) \right] \pi(ds|s\_) \tag{28}$$

where maximization is subject to $Z(s) \leq \overline{Z},\ B(s) \in [\underline{B}, \overline{B}]$ and

$$B(s) = R(s, s\_)B\_ + g(s) - \Theta(s)Z(s) \text{ for all } s. \tag{29}$$

In the interior, optimal debt satisfies

$$V\left(\tilde{B}_t, s_t\right) = \mathbb{E}_t V\left(\tilde{B}_{t+1}, s_{t+1}\right) + \beta \text{cov}_t(R_{t+1}, V'(\tilde{B}_{t+1}, s_{t+1})), \tag{30}$$

which extends the martingale equation (15) to persistent shocks.

As a counterpart to expression (16), we now define the risk-minimizing government debt B^* for the general case being studied here. As before, the Ramsey planner chooses government debt to minimize risk and fluctuations in the tax rate. The shocks here introduce additional considerations not present in the Section 2 baseline model. First, fluctuations in productivity imply that tax revenues are stochastic even when the tax rate is constant. Fix the tax rate at level τ and observe that primary deficit X_τ, defined as the difference between expenditures and tax revenues, is

$$X_\tau(s) = g(s) - \Theta(s)(1-\tau)^{\frac{1}{\gamma}}\tau. \tag{31}$$

Fluctuations in the primary government deficit are driven by shocks to both government expenditures and to productivity. Furthermore, when these processes are persistent, the current state s_t conveys information about future primary deficits. Now government debt will play an important role in hedging fluctuations *in the expected present value of primary deficits.*

For a random variable $x(s)$ that is a function of the current state only, a discounted present value of x conditional on s is $PV(x;s) \equiv \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t x_t | s_0 = s\right]$. Since the planner keeps the tax rate approximately constant, the mean of the invariant distribution for debt and the level of tax rate are linked through the government budget constraint by

$$\left(\frac{1-\beta}{\beta}\right)B = \bar{g} - \bar{\Theta}(1-\tau)^{\frac{1}{\gamma}}\tau, \tag{32}$$

which defines an implicit function $\tau(B)$. We define B^* as the level of debt that minimizes fluctuations in $PV(X_{\tau(B)};s)$:

$$B^* \equiv \arg\min_B \operatorname{var}\left[RB + PV(X_{\tau(B)})\right]. \tag{33}$$

We define Z^* as $Z^* \equiv \frac{1}{\bar{\Theta}}\left[\bar{g} + \frac{1-\beta}{\beta}B^*\right]$.

We again use a second order approximation of policies to show that B^* is the long-run target level of government debt. To state things compactly, it helps to define two mappings. For a pair of random variables $x(s,s\_)$, $y(s,s\_)$, the covariance conditional on $s\_$ is

$$\mathcal{C}^{x,y}(s_-) \equiv \int x(s,s\_)y(s_-,s)\pi(ds|s_-) - \left(\int x(s,s\_)\pi(ds|s_-)\right)\left(\int y(s,s\_)\pi(ds|s_-)\right)$$

and the conditional mean of $x(s)$ is

$$\mathrm{E}(x;s\_) \equiv \int x(s,s\_)\pi(ds|s_-).$$

Note that both $\mathcal{C}^{x,y}(\cdot)$ and $\mathrm{E}(x;\cdot)$ are random variables on S. Taking a Taylor expansion of optimal policies that attain the optimal value function $V(B\_,s\_)$ that satisfies Bellman equation (28) along lines taken in Section 2 we get[17]

$$\begin{aligned}\tilde{B}(s,B\_,s\_) &= B\_ + [g(s) - (1-\beta)PV(g;s)] - \bar{\Phi}(B\_)[\Theta(s) - (1-\beta)PV(\Theta;s)] + B\_[\beta R(s,s\_) - 1] \\ &- (1-\beta)\beta^2\left[B\_PV\left(\mathcal{C}^{R,R};s\right) + PV\left(\mathcal{C}^{R,PV(g)};s\right) - \bar{\Phi}(B\_)PV\left(\mathcal{C}^{R,PV(\Theta)};s\right)\right] + \mathcal{O}(\sigma^3,(1-\beta)\sigma^2),\end{aligned} \tag{34}$$

where $\bar{\Phi}(B\_) = \frac{1}{\bar{\Theta}}\left[\left(\frac{1-\beta}{\beta}\right)B\_ + \bar{g}\right]$.

The first line on the right side of equation (34) collects first order expansion terms that capture direct effects of shocks to g,p,θ on the asset positions. Government debt increases if the current realization of g is greater than annuitized expected future expenditures, $(1-\beta)PV(g;s)$; if current realization of productivity is less than then annuitized expected future productivity

[17] A detailed derivation of equation (34) appears in Appendix A.4.

$(1-\beta)\,PV(\Theta;s)$; or if the interest payments on debt are unexpectedly high. These terms express how optimal policy uses debt to smooth aggregate shocks and embody principles conveyed by Barro (1979). The second line on the right side of equation (34) collects second order terms that consist of conditional variances and covariances of the return with expenditure and productivity shocks. They capture hedging motives.

It is convenient to re-write equation (34) in terms of ergodic moments of (g, p, θ). For a random variable $x(s, s\_)$, let $\mathbb{E}x = \int\int x(s, s\_)\pi(ds|s\_)\lambda(ds\_)$ be its ergodic mean. Similarly, let $\mathrm{var}\,(x)$ and $\mathrm{cov}\,(x, y)$ denote ergodic variances and covariances of random variables x and y, respectively. In Appendix A.4, we show that under our assumption about π, we can write (34) as

$$\tilde{B}\,(s, B\_, s\_) = B\_ + [g(s) - (1-\beta)\,PV(g;s)] - \bar{\Phi}(B\_)\,[\Theta(s) - (1-\beta)\,PV(\Theta;s)] + B\_\,[\beta R(s, s\_) - 1]$$
$$- \beta^2\left[B\_\mathrm{var}(R) + \mathrm{cov}(R, PV(g)) - \bar{\Phi}(B\_)\mathrm{cov}(R, PV(\Theta))\right] + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2). \tag{35}$$

For a random variable $x(s, s\_)$, let $\hat{x}(s, s\_) \equiv x(s, s\_) - \mathrm{E}(x; s\_)$. We can use (35) to obtain

Proposition 6. *The invariant distribution of* $\left\{\tilde{B}_t, \tilde{Z}_t\right\}_t$ *has*

- **Means**
$$\mathbb{E}\left(\tilde{B}_t\right) = B^* + \mathcal{O}(\sigma, (1-\beta)), \quad \mathbb{E}\left(\tilde{Z}_t\right) = Z^* + \mathcal{O}(\sigma, (1-\beta)),$$

- **Speeds of reversion to means**
$$\frac{\mathbb{E}_t\left(\tilde{B}_{t+1} - B^*\right)}{\tilde{B}_t - B^*} = \frac{\mathbb{E}_0\left(\tilde{Z}_{t+1} - Z^*\right)}{\mathbb{E}_0\left(\tilde{Z}_t - Z^*\right)} = \frac{1}{1 + \beta^2\mathrm{var}\,(R)} + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2).$$

- **Variances** Define $\mathbb{B}(s) \equiv B^* - \beta\left[\mathrm{E}\,(PV(g-\bar{g}; s'); s) - \bar{\Phi}(B^*)\mathrm{E}\left(PV(\Theta - \bar{\Theta}; s'); s\right)\right]$. Then
$$\mathrm{var}\left(\tilde{B}_t - \mathbb{B}_t\right) = \mathrm{var}\left(\hat{PV}(g) - \bar{\Phi}(B^*)\hat{PV}(\Theta) + \mathbb{B}\hat{R}\right) + \mathcal{O}(\sigma, (1-\beta)),$$
$$\mathrm{var}\left(\tilde{Z}_t\right) = \left(\frac{1-\beta}{\beta}\right)^2 \mathrm{var}\left(\tilde{B}_t - \mathbb{B}_t\right) + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2).$$

Proposition 6 shows that just as in Section 2, the Ramsey planner chooses government debt to minimize risk and keep the tax rate approximately constant. One can extend our approximations (18) to show that the Euler equation (30) induces reversion of government debt to a risk-minimizing level. Productivity shocks now induce fluctuations in tax revenues even when the tax rate is constant.

The risk-minimizing debt level B^* can be computed from formula (33) and further simplified after we observe that $\frac{\partial}{\partial B}\tau(B) = \mathcal{O}(1-\beta)$. Given this, we have

$$\begin{aligned}
B^* &= -\frac{\operatorname{cov}\left(R, PV(X_{\tau(B)})\right)}{\operatorname{var}(R)} + \mathcal{O}(\sigma,(1-\beta)) \text{ for any } B \\
&= -\frac{\operatorname{cov}\left(R, PV(g)\right) - \frac{\bar{g}}{\bar{\Theta}}\operatorname{cov}\left(R, PV(\Theta)\right)}{\operatorname{var}(R)} + \mathcal{O}(\sigma,(1-\beta)). \qquad (36)
\end{aligned}$$

The simple formula (36) for the approximate risk-minimizing debt level presents a further insight that we shall exploit in Sections 3.4 and 4. It shows that the endogenous covariances that appear in this formulas are not very sensitive to values of $\tau(B)$ at which they are evaluated. That means tat if we were to observe data generated under a suboptimal tax rate policy $\tau(B)$, observations of the primary deficit $X_{\tau(B)}$ would still allow us to compute the optimal level of debt B^* accurately by using (36).

We end this section by applying our formulas when g, Θ, p obey the AR(1) processes

$$g_t = (1-\rho_g)\bar{g} + \rho_g g_{t-1} + \varepsilon_{g,t},$$

$$\Theta_t = (1-\rho_\Theta)\bar{\Theta} + \rho_\Theta \Theta_{t-1} + \varepsilon_{\Theta,t}$$

$$p_t = \bar{p} + \varepsilon_{p,t},$$

where $\varepsilon_{g,t}, \varepsilon_{p,t}, \epsilon_{\Theta,t}$ are i.i.d over time with zero means. Now $\operatorname{cov}(R, PV(g)) = \frac{\operatorname{cov}(R,g)}{1-\rho_g\beta}$ and $\operatorname{cov}(R, PV(\Theta)) = \frac{\operatorname{cov}(R,\Theta)}{1-\rho_\Theta\beta}$. Therefore

$$\begin{aligned}
B^* &= -\left(\frac{1}{1-\rho_g\beta}\right)\frac{\operatorname{cov}(R,g)}{\operatorname{var}(R)} + \frac{\bar{g}}{\bar{\Theta}}\left(\frac{1}{1-\rho_\Theta\beta}\right)\frac{\operatorname{cov}(R,\Theta)}{\operatorname{var}(R)}, \\
&= -\left(\frac{\beta}{1-\rho_g\beta}\right)\frac{\operatorname{cov}(\epsilon_p,\epsilon_g)}{\operatorname{var}(\epsilon_p)} + \frac{\bar{g}}{\bar{\Theta}}\left(\frac{\beta}{1-\rho_\Theta\beta}\right)\frac{\operatorname{cov}(\epsilon_p,\epsilon_\Theta)}{\operatorname{var}(\epsilon_p)}.
\end{aligned}$$

This formula shows how autocorrelations affect the target level of government debt. For instance, keeping ρ_Θ fixed, higher persistence of the expenditure shocks as measured by ρ_g implies a higher absolute value of government debt asymptotically. The sign of the covariance between returns and the primary government deficit determines the sign of the mean level of government debt.

3.4 Risk aversion and endogenous returns

We extend our analysis to a setting in which the representative agent has preferences that display risk-aversion. We retain other assumptions of Section 3.3 but now allow curvature in the utility of consumption by assuming that preferences are described by

$$U(c,l) = \frac{c^{1-\alpha}-1}{1-\alpha} - \frac{l^{1+\gamma}}{1+\gamma}. \qquad (38)$$

A constant elasticity of substitution assumption simplifies exposition, but our results prevail for U's that are strictly concave in $(c,-l)$ and twice continuously differentiable. We let $U_{x,t}$ or $U_{xy,t}$ denote first and second derivatives of U with respect to $x, y \in \{c, l\}$. We assume that natural debt limits restrict the consumer, which ensures that first-order conditions are satisfied off corners.

An allocation $\{c_t, l_t, \mathrm{B}_t\}_t$ is a competitive equilibrium if and only if it satisfies the feasibility constraint (27) and implementability conditions

$$U_{c,t}\mathrm{B}_t + U_{c,t}\left[\theta_t l_t + \frac{U_{l,t}}{U_{c,t}} l_t - g_t\right] = \frac{p_t U_{c,t}}{\beta \mathbb{E}_{t-1} p_t U_{c,t}} U_{c,t-1}\mathrm{B}_{t-1} \quad t \geq 1, \tag{39}$$

$$c_0 + b_0 = -\frac{U_{l,0}}{U_{c,0}} l_0 + p_0 \beta^{-1} \mathrm{B}_{-1}. \tag{40}$$

An optimal allocation maximizes $\mathbb{E}_0 \sum_t \beta^t U(c_t, l_t)$ subject to constraints (27), (39), and (40).

It is helpful to redefine variables. Let $\mathcal{B}_t \equiv U_{c,t}\mathrm{B}_t$, $\mathcal{R}_t \equiv \frac{U_{c,t} p_t}{\beta \mathbb{E}_{t-1} U_{c,t} p_t}$, and $\mathcal{X}_t \equiv U_{c,t}\left[g_t - \tau_t \theta_t l_t\right]$ be marginal utility adjusted debt, return and primary deficit. Using the household's first-order necessary conditions and the resource constraint, at any state s for a given tax rate τ, a household's consumption $c_\tau(s)$ satisfies

$$(1-\tau)\,\theta(s) c_\tau(s)^{-\alpha} + \left(\frac{c_\tau(s) + g(s)}{\theta(s)}\right)^{\gamma} = 0. \tag{41}$$

Along any history $\left(s^{t-1}, s_t\right)$ effective returns and effective deficits can be expressed in terms of exogenous states s_t and a period-t tax rate τ as

$$\begin{aligned} \mathcal{R}_\tau\left(s_t, s^{t-1}\right) &= \frac{c_\tau\left(s_t\right)^{-\alpha} p\left(s_t\right)}{\beta \int c_\tau\left(s'\right)^{-\alpha} p\left(s'\right) \pi\left(ds' | s_{t-1}\right)}, \\ \mathcal{X}_\tau\left(s_t, s^{t-1}\right) &= \left(\frac{c_\tau\left(s_t\right) + g\left(s_t\right)}{\theta\left(s_t\right)}\right)^{1+\gamma} - c_\tau\left(s_t\right)^{1-\alpha}. \end{aligned}$$

These transformations allow us to assert that the Ramsey planner's optimal value function for $t \geq 1$ satisfies the Bellman equation:

$$V\left(\mathcal{B}\_, s\_\right) = \max_{\tau(\cdot), \mathcal{B}(\cdot)} \int \left[U\left(c_{\tau(s)}(s), \frac{c_{\tau(s)}(s) + g(s)}{\theta(s)}\right) + \beta V\left(\mathcal{B}(s), s\right)\right] \pi\left(ds | s\_\right) \tag{42}$$

where maximization is subject to

$$\mathcal{B}(s) = \mathcal{R}_{\tau(s)}\left(s, s\_\right) \mathcal{B}\_ + \mathcal{X}_{\tau(s)}\left(s\right) \text{ for all } s. \tag{43}$$

Problem (42) closely resembles problem (28) except that all variables have been transformed into their effective counterparts.[18] The planner now uses effective debt to smooth risk, and the evolution of optimal effective government debt level satisfies

$$V'\left(\tilde{\mathcal{B}}_t, s_t\right) = \mathbb{E}_t V'\left(\tilde{\mathcal{B}}_{t+1}, s_{t+1}\right) + \beta \mathrm{cov}_t(\mathcal{R}_{t+1}, V'(\tilde{\mathcal{B}}_{t+1}, s_{t+1})),$$

[18] The planner's problem at $t=0$ at initial debt B_{-1} and state s_{-1} is

$$\max_{\tau(\cdot), \mathcal{B}(\cdot)} \int \left[U\left(c_{\tau(s)}(s), \frac{c_{\tau(s)}(s) + g(s)}{\theta(s)}\right) + \beta V\left(\mathcal{B}(s), s\right)\right] \pi\left(ds | s_{-1}\right)$$

subject to

$$\mathcal{B}(s_0) = \mathcal{X}_\tau\left(s\right) + U_c\left(c_{\tau(s)}(s)\right) p(s) \beta^{-1} \mathrm{B}_{-1} \quad \forall s$$

an analogue of (30). The economic intuition for this equation is that the planner still uses covariance of returns with shadow cost of debt to hedge risk, but adjusts all the variables for the shadow costs of raising revenues.

We can use insights from Section 3.3 to define a risk-minimizing level of effective debt as

$$\mathcal{B}^* \equiv \arg\min_{\mathcal{B}} \operatorname{var}\left[\mathcal{R}\mathcal{B} + PV(\mathcal{X}_{\tau(\mathcal{B})})\right], \tag{44}$$

where $\tau(\mathcal{B})$ satisfies the following ergodic version of the government budget constraint

$$\left(\frac{1-\beta}{\beta}\right)\mathcal{B} = \mathbb{E}\mathcal{X}_\tau(\cdot). \tag{45}$$

We extend Proposition 6 to accommodate risk averse preferences.

Proposition 7. *The ergodic mean and the speed of mean reversion of effective debt* $\left\{\tilde{\mathcal{B}}_t\right\}_t$ *are*

$$\mathbb{E}\tilde{\mathcal{B}}_t = \mathcal{B}^* + \mathcal{O}(\sigma, 1-\beta),$$

$$\frac{\mathbb{E}_t\left(\tilde{\mathcal{B}}_{t+1} - \mathcal{B}^*\right)}{\tilde{\mathcal{B}}_t - \mathcal{B}^*} = \frac{1}{1+\beta^2\operatorname{var}(\mathcal{R}_{\tau(\mathcal{B}^*)})} + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2).$$

Furthermore, $\mathcal{B}^*$ *satisfies*

$$\mathcal{B}^* = -\frac{\operatorname{cov}\left(\mathcal{R}_{\tau(\mathcal{B})}, PV\left(\mathcal{X}_{\tau(\mathcal{B})}\right)\right)}{\operatorname{var}\left(\mathcal{R}_{\tau(\mathcal{B})}\right)} + \mathcal{O}(1-\beta) \textit{ for all } \mathcal{B}. \tag{46}$$

Proposition 7 confirms our theme that a target level of government debt under the optimal plan solves a variance-minimization problem. It also extends a finding from equation (36) that while second moments of returns and the primary government deficit depend on government policy, effects of the tax rate are is small, so omitting them lead to errors of order only $\mathcal{O}(1-\beta)$. This means that we can also simply estimate the variance of effective returns and their covariance with the effective deficit directly and then use that estimate to estimate a risk-minimizing level of government debt. We apply that procedure in Section 4.

Formula (46) has implications about an optimal level of risk-free debt that relate to findings of Aiyagari et al. (2002). The return on risk-free debt is known one period in advance, but the effective return are not. In particular, the effective return is high in states in which consumption is low, namely, states in which the primary government deficit is high, either because government expenditures are high or productivity is low, making $\operatorname{cov}(\mathcal{R}_\tau, PV(\mathcal{X}_\tau)) > 0$. Therefore, formula (46) implies that an optimal long-run level of risk-free debt is negative, i.e., the planner accumulates assets. Furthermore, since aggregate consumption growth is not volatile, at least in U.S. data, $\operatorname{var}(\mathcal{R}_\tau)$ would be low in most U.S. calibrations, implying that the long-run asset level should be quite high (see also Example 2 in Section 3.2). This provides intuition for some of the numerical findings in Aiyagari et al. (2002) and some subsequent contributions too.

We can apply insights from Section 3.2 to situations in which the planner manages a portfolio of K securities. A version of the planner's Bellman equation (42), modified to have effective total assets to be the state variable, extends along the lines in (22). In the interior, a martingale equation restricts every security, namely,

$$V'\left(\tilde{\mathcal{B}}_t, s_t\right) = \mathbb{E}_t \mathcal{R}^k_{t+1} V'\left(\tilde{\mathcal{B}}_{t+1}, s_{t+1}\right) = \mathbb{E}_t V'\left(\tilde{\mathcal{B}}_{t+1}, s_{t+1}\right) + \beta \text{cov}_t(\mathcal{R}^k_{t+1}, V'(\tilde{\mathcal{B}}_{t+1}, s_{t+1})), \quad (47)$$

where $\mathcal{R}^k_{t+1}$ is the effective return on security k. Equation (47) implies equation (34) of Farhi (2010) that describes CCAPM Euler equations. Now let $\mathcal{R}^k_\tau$ be the effective returns on asset k evaluated at tax rate τ and let $\mathbf{R}_\tau = \left[\mathcal{R}^1_\tau ... \mathcal{R}^K_\tau\right]$ be a matrix of these returns. Combining the insights from this section and Lemma 4, it follows that the risk-minimizing portfolio can be approximated, up to the order $\mathcal{O}(\sigma, (1-\beta))$, by

$$-\mathbb{C}[\mathbf{R}_{\tau(0)}, \mathbf{R}_{\tau(0)}]^{-1} \mathbb{C}[\mathbf{R}_{\tau(0)}, PV(\mathcal{X}_{\tau(0)})]. \quad (48)$$

4 A quantitative example

We now study an economy with a risk-averse representative consumer together with g, p, θ processes calibrated to match stylized U.S. business cycle facts during the post WWII period. We use our expressions from Proposition 7 and the equation (48) risk-minimizing portfolio and related extensions of expressions for other moments reported in Proposition 6. Among other things, we use these calculations to verify the accuracy of our approximations for the ergodic behavior of government debt, the tax rate, and tax collections under an optimal plan.

We set utility function parameters α, γ, β equal to $1, 2, 0.98$. We begin by assuming that households and the government trade a single one-period security and parameterize a stochastic process for (θ_t, p_t, g_t) in terms of the following AR(1) specifications:

$$\begin{aligned}
\ln \theta_t &= \rho_\theta \ln \theta_{t-1} + \sigma_\theta \epsilon_{\theta,t} \\
\ln g_t &= \ln \bar{g} + \chi_g \epsilon_{\theta,t} + \sigma_g \epsilon_{g,t} \\
\ln p_t &= \chi_p \epsilon_{\theta,t} + \sigma_p \epsilon_{p,t},
\end{aligned}$$

where $\epsilon_{\theta,t}$, $\epsilon_{g,t}$ and $\epsilon_{p,t}$ are i.i.d. standard normal random variables.

Our parameterizations of productivity and government expenditures are standard, but our calibration of asset payoff is less common. The literature typically assumes that the real payoff on government debt is risk-free and calculates returns on that asset from a marginal utility of a representative consumer within a neoclassical growth model. This approach unfortunately implies asset returns that are not consistent with observed returns on government debt. That deficiency matters for us because our formulas assert that the variance and covariance of returns on government debt are important determinants of optimal debt management. Therefore, we simply set parameters of the stochastic process of payoffs p_t to assure that the return on the government's portfolio matches the return on the security held or issued by the government in our model.[19]

Table II documents our calibration targets for parameters $(\bar{g}, \rho_\theta, \chi_g, \chi_p, \sigma_\theta, \sigma_g, \sigma_p)$ in terms

[19]A more sophisticated approach would be to model the reason for the fluctuations in real returns on government debt explicitly. The asset pricing literature in finance proposes several ways to do this. One approach is the discount factor shock model of Albuquerque et al. (2016 forthcoming). Appendix B shows that our model with payoff shocks to debt is essentially equivalent to a model with discount factor shocks but risk-free debt.

of moments of output, government expenditures, and bond returns. We use time series for these variables for 1947-2014 at annual frequencies. Except for returns, we took logarithms of all variables and then Hodrick-Prescott pre-filtered them, using a smoothing parameter equal to 6.25. For output y_t and government expenditures g_t, we use Bureau of Economic Analysis data for aggregate real labor earnings and federal government consumption expenditures plus transfer payments.[20] We measure B_t as the real market value of gross federal debt series published by the Federal Reserve Bank of Dallas.[21]

We propose two measures of returns on government debt. As a baseline, we impute real returns R_t using data on the real federal primary deficit[22] X_t and market value of government debt B_t. The observed duration of government debt has been approximately constant, allowing us to write the government budget constraint as

$$(p_t + q_t)\mathrm{B}_{t-1} = q_t \mathrm{B}_t - X_t. \tag{49}$$

Multiply and divide the first term by q_{t-1} and use the fact that the holding period return for long term debt is $R_t = \frac{q_t + p_t}{q_{t-1}}$ to rewrite equation (49) as

$$R_t = \frac{B_t - X_t}{B_{t-1}}, \tag{50}$$

where $B_t = q_t \mathrm{B}_t$ is the observed market value of government debt. The average annual return in our sample is about 5% and its standard deviation is 5%. As an alternative measure, we also calibrate payoffs to match the moments of the real one year U.S. treasury yield, obtained from release H.15 of the Board of Governors of the Federal Reserve System. The average return in our sample is 2.0% with a standard deviation of 2.6%. The main difference between the two return measures comes from the fact that we capital gains from revaluations of long term debt are captured in imputed returns, but not in one year treasury yields.

Returns is in our model are endogenous and depend both on parameters and on government policy $\{\tau_t, B_t\}_t$. We assume that the tax rate conformed to the rule

$$\tau_t = (1-\rho_\tau)\bar{\tau} + \rho_\tau \tau_{t-1} + \rho_Y \log y_t + \rho_{Y\_} \log y_{t-1} + \rho_g g_t + \rho_{g\_} g_{t-1} + \rho_R R_t + \rho_{R\_} R_{t-1} + \rho_{B\_} \log B_t, \tag{51}$$

whose coefficients we estimated with an OLS regression using our series on output, expenditure, returns, debt, and an average marginal income tax rate τ_t obtained from Barro and Redlick (2011). Our specification (51) is flexible enough to capture how tax rates are persistent and how they adjust to movements in government expenditures, returns, and the level of government debt. We report estimated coefficients of (51) in Table III and the in-sample fit in Figure II. Given our estimated tax rule, we set debt $\{B_t\}_t$ to satisfy the government's budget constraint. Appendix C provides details about how we compute a competitive equilibrium given government

[20]Since in our model we abstract from capital, our measure of output y is aggregate labor earnings. Results remain essentially unchanged if we use GDP per capita instead.

[21]Calculation of this series takes into account outstanding marketable and non-marketable debt of different maturities issued by the Treasury and uses current market prices to convert par value to market value.

[22]We measure this as government expenditure, i.e., federal consumption and transfer payments, minus total federal tax receipts, both from the Bureau of Economic Analysis.

policy $\{\tau_t, B_t\}_t$. Table II summarizes parameters values and the fit of a competitive equilibrium outcomes to U.S. data.

Using this calibration, we compute a global approximation to the Ramsey allocation. Appendix C reports details about the numerical procedure. In Table IV, we compare predictions of our quadratic approximations about the behavior of government debt and tax revenues to those obtained by using a more accurate global numerical procedure. Given our assumption of logarithmic utility, effective debt and returns are simply $\mathcal{B}_t = B_t/c_t$ and $\mathcal{R}_t = R_t c_t/c_{t-1}$. Following Proposition 7, we use equation (46) evaluated at $\tau(0)$ to calculate the risk-minimizing level of debt and $\text{var}(\mathcal{R}_{\tau(0)})$ to compute the speed of mean reversion. We similarly use equations in Proposition 6, now written in terms of effective units, to compute the ergodic variance of effective debt and moments of tax rates $Z_t = \tau_t (1-\tau_t)^{\frac{1}{\gamma}}$.

We computed the ergodic distribution by simulating policies computed using the global approximation method. The first two columns in Table IV show that for the baseline calibration, our expressions for the ergodic distribution of debt and tax revenues approximate well those obtained from the simulations. As an illustration of how the approximations do away from the ergodic distribution, we plot $\mathbb{E}_0 \mathcal{B}_t$ using

$$\mathbb{E}_0[\mathcal{B}_t - \mathcal{B}^*] \approx (\mathcal{B}_t - \mathcal{B}^*) \left(\frac{1}{1+\beta^2 \text{var}(\mathcal{R}_{\tau(0)})} \right)^t, \tag{52}$$

and compare it to the mean path constructed using 10000 simulations under policies computed using the more accurate global methods of length 15000 periods. Figure III indicates that formula (52) gives a very accurate approximation for the entire path and not just its long run target level of effective debt.

An insight of Proposition 7 is that covariances and variances are not very sensitive to the policies under which they are evaluated. Therefore, one should expect that the calculation of these variances in the data, generated by the actual rather than an optimal policy, produce reliable estimates of the optimal long-run debt level. We verify this as follows. Consider a simple first order VAR

$$\begin{bmatrix} \mathcal{X}_t \\ \log y_t \end{bmatrix} = A \begin{bmatrix} \mathcal{X}_{t-1} \\ \log y_{t-1} \end{bmatrix} + \Sigma \begin{bmatrix} \epsilon_{\mathcal{X},t} \\ \epsilon_{y,t} \end{bmatrix}.$$

Let $[a_{\mathcal{X}} \;\; a_y]$ be the first row of the matrix $[I-\beta A]^{-1}$. Then the expected present value of the primary government surplus conditional on $(\mathcal{X}_t, \log y_t)$ is

$$PV\left(\mathcal{X}; (\mathcal{X}_t, \log y_t)\right) = [a_{\mathcal{X}} \;\; a_y] \begin{bmatrix} \mathcal{X}_t \\ \log y_t \end{bmatrix},$$

and an appropriate estimate of the target level of debt is

$$\mathcal{B}^* = -a_y \frac{\text{cov}(\mathcal{R}_t, \log y_t)}{\text{var}(\mathcal{R}_t)} - a_{\mathcal{X}} \frac{\text{cov}(\mathcal{R}_t, \mathcal{X}_t)}{\text{var}(\mathcal{R}_t)}.$$

We use our time series for returns, consumption, output, and the primary government deficit to construct time series of $\mathcal{R}_t$ and $\mathcal{X}_t$. Table V presents the estimated coefficients that we then use to estimate both the target level of effective debt and the speed of the mean reversion reported in the column titled "VAR" in Table IV.

The findings in Table IV convey that at our baseline calibration of the long-run effective debt is close to zero, that the convergence is slow (half life of 250 years), that government debt has large fluctuations (the standard deviation is 20%), while there are small movements in the tax rate and tax revenues (whose standard deviations is 0.5%). A key empirical fact that drives these results is that a substantial component of fluctuations in returns is uncorrelated with fundamentals. That makes holding large positions frustrate the hedging motive and drives the optimal plan towards low assets.

In our baseline, we chose the payoff process to match imputed returns on the total debt portfolio traded by the government. To check robustness of our results we show that our approximation procedure continues to work when we instead measure the returns using the 1 year U.S. treasury yield or assume that the real debt traded by the government is risk-free. In Table IV, the columns "1 yr. yield" and "risk-free debt" report moments of the ergodic distribution for our calibrated economy in which we set $\chi_p = -0.10, \quad \sigma_p = 0.02$ to match the standard deviation of 1 year U.S. treasury yields and the correlation of those yields with output, which are 2.6% and -0.20 in our sample, respectively, and then, alternatively, $\chi_p = \sigma_p = 0$ to obtain the risk free payoff. These alternative assumptions about returns progressively weaken the orthogonal component. Consistent with our discussion of equation (46), the government holds a larger asset position (i.e., a negative debt) to exploit the stronger positive correlation of returns and deficits. Because the speed of mean reversion is inversely related to the volatility of returns, the half-life of debt increases from 237 years in the baseline calibration to 655 years for the calibration with one year yields and increases further to 1244 years for the risk-free debt. In all of these settings, our simple formulas capture the comparative outcomes extremely well.

We now extend our analysis to allow the government to trade multiple assets. We pursue two aims with this extension. First, we want to evaluate the accuracy of approximations provided by equation (48). Second, we want to highlight additional insights about optimal government portfolio management and to re-examine an argument of Lucas and Zeldes (2009) that it is optimal for a government to take a positive position in a risky security that pays a risk premium. Although our problem has some features in common with the problem solved by Merton (1969), there are two critical differences: our problem is posed within a general equilibrium in which a Ramsey planner takes into account how its government actions affect asset returns; and the Ramsey planner is benevolent.

We fix parameters as described above except that now we assume that the government trades two securities. One is a riskless real bond; the logarithm of the payoff on the other security is described by $\ln p_t = \chi_p \epsilon_{\theta,t} + \sigma_p \epsilon_{p,t}$, where χ_p, σ_p are now calibrated to match the correlation of dividends on the S&P500 with output and the standard deviation of these dividends, which for our sample take the values 0.30 and 4.5%. Making the payoffs positively comove with TFP makes this asset risky resulting in higher expected holding period returns relative to the risk free rate.[23] We set the initial debt at 130% of output.

[23]Given our assumption of isoelastic preferences, we cannot match the magnitude of the risk premium quan-

From Table VI we see that the formula in equation (48) accurately describes the long-run portfolio in addition to the speed of mean reversion and standard deviation of total assets. In the long run, the optimal plan has negative debt invested almost solely in the risk free asset. Initially, when it is indebted, the government shorts the stock market. Although the initial short position in the stock market exposes the government to the orthogonal component $\epsilon_{p,t}$ in the payoff, temporarily it provides a good hedge by delivering higher returns in times of low TFP. Eventually, the government uses only the risk free bond to hedge. The dynamics of portfolio re-balancing are consistent with Example 3 from Section 3.2. The two lines in Figure IV show the marginal utility adjusted government positions in the risk-free security (solid line) and in the risky security (dashed line).

Our quantitative analysis confirms the optimality of a portfolio management rule based on the variance-minimization principle outlined in Section 3.2 and cautions against following recommendations that a government should on the margin invest in assets that pay a risk premium. In our economy, the Ramsey planner shares households' aversion to consumption risk, an aversion that in general equilibrium requires a return premium to compensate for bearing risk. The Ramsey planner finds it optimal to invest in such assets only in so far as doing so helps to reduce the total riskiness of gross government expenditures.

5 Concluding remarks

This paper characterizes optimal debt management and flat rate taxation in a fairly general incomplete markets model. We express dynamic hedging motives in a terms of a fiscal risk minimization problem. We present simple formulas for the mean, variance, and speed of convergence to an ergodic distribution of government debt. We analyze some extensions of our basic environment, an endeavor we pursue more in Bhandari et al. (2015b), where we study economies whose substantial *ex ante* heterogeneity coming from persistent differences in skills that unleash social motives for redistribution and social insurance. The analysis here sets the stage for such extensions – partly by providing appropriate tools for approximating equilibria well and for formulating Ramsey problems in mathematically convenient ways, and partly by isolating transcendent forces that ultimately determine transient and long-run dynamics of government debt and taxes in richer settings. For example, appropriately adjusted fiscal risk minimization problems continue to shape the mean of an ergodic distribution of government debt, while the hedging costs of being away from that fiscal risk-minimizing debt level shape speeds of convergence. Another extension, Bhandari et al. (2015a), uses the empirical properties of returns across maturities to compute an optimal maturity structure of government debt.

titatively but we conjecture that our approach extends to the Epstein-Zin preferences and richer environments with more realistic implications for asset pricing behavior such as the one considered by Albuquerque et al. (2016 forthcoming). We leave this extension it to future work.

References

Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala. 2002. "Optimal Taxation without State-Contingent Debt." *Journal of Political Economy*, 110(6): 1220–1254.

Albuquerque, Rui, Martin Eichenbaum, Victor Luo, and Sergio Rebelo. 2016 forthcoming. "Valuation Risk and Asset Pricing." *Journal of Finance*.

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

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy*, 87(5): 940–971.

Barro, Robert J. 1999. "Notes on Optimal Debt Management." *Journal of Applied Economics*, 2(2): 281–289.

Barro, Robert J. 2003. "Optimal Management of Indexed and Nominal Debt." *Annals of Economics and Finance*, 4 1–15.

Barro, Robert J., and Charles J. Redlick. 2011. "Macroeconomic Effects From Government Purchases and Taxes." *Quarterly Journal of Economics*, 126(1): 51–102.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas Sargent. 2015a. "The Optimal Term Structure of Government Debt."

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas Sargent. 2015b. "Taxes, Debt and Redistribution with Incomplete Markets."

Bohn, Henning. 1990. "Tax Smoothing with Financial Instruments." *The American Economic Review*, 80(5): 1217–1230.

Buera, Francisco, and Juan Pablo Nicolini. 2004. "Optimal Maturity of Government Debt without State Contingent Bonds." *Journal of Monetary Economics*, 51(3): 531–554.

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.

Debortoli, Davide, Ricardo Nunes, and Pierre Yared. 2016 forthcoming. "Optimal Time-Consistent Government Debt Maturity." *Quaterly Journal of Economic*.

Dunford, Nelson, and Jacob T Schwartz. 1963, 1966. *Linear Operators, Part I General Theory and Part II Spectral Theory.*: Interscience Publishers.

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.

Fleming, Wendell H. 1971. "Stochastic Control for Small Noise Intensities." *SIAM Journal on Control*, 9(3): 473–517.

Hopenhayn, Hugo A., and Edward C. Prescott. 1992. "Stochastic Monotonicity and Stationary Distributions for Dynamic Economies." *Econometrica*, 60(6): 1387–1406.

Kopecky, Karen, and Richard Suen. 2010. "Finite State Markov-chain Approximations to Highly Persistent Processes." *Review of Economic Dynamics*, 13(3): 701–714.

Lucas, Deborah J, and Stephen P Zeldes. 2009. "Pensions and Health Care: Fiscal Challenges for State and Local Governments-How Should Public Pension Plans Invest?" *American Economic Review*, 99(2): 527–532.

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.

Lustig, Hanno, Christopher Sleet, and Sevin Yeltekin. 2008. "Fiscal Hedging with Nominal Assets." *Journal of Monetary Economics*, 55(4): 710 – 727.

Markowitz, Harry. 1952. "Portfolio Selection." *The Journal of Finance*, 7(1): 77–91.

Merton, Robert C. 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case." *The Review of Economics and Statistics*, 51(3): 247–257.

Quah, John K.-H. 2007. "The Comparative Statics of Constrained Optimization Problems." *Econometrica*, 75(2): 401–431.

Schmitt-Grohe, Stephanie, and Martin Uribe. 2004. "Solving dynamic general equilibrium models using a second-order approximation to the policy function." *Journal of Economic Dynamics and Control*, 28(4): 755–775.

Siu, Henry E. 2004. "Optimal Fiscal and Monetary Policy with Sticky Prices." *Journal of Monetary Economics*, 51(3): 575–607.

A Proofs

A.1 Proofs for Section 2

We first show preliminary results that we discussed in Section 2. Let $\left\{\tilde{B}_t, \tilde{l}_t\right\}_t$ be a solution to (10) and $\tilde{Z}_t = \tilde{l}_t - \tilde{l}_t^{1+\gamma}$.

Lemma 5. *$\tilde{l}_t^{\gamma} \geq \frac{1}{1+\gamma}$ for all t and there is one to one map between $\tilde{l}_t$ and $\tilde{Z}_t$ and $\tilde{\tau}_t$ and $\tilde{Z}_t$ with $\tilde{Z}_t \leq \overline{Z}$ for all t. The function $\Psi : (-\infty, \overline{Z}] \to \mathbb{R}$ is strictly decreasing, strictly concave, differentiable and satisfies $\lim_{Z\to-\infty} \Psi'(Z) = 0$ and $\lim_{Z\to\overline{Z}} \Psi'(Z) = -\infty$.*

Proof. First we show that $\tilde{l}_t^{\gamma} \geq \frac{1}{1+\gamma}$ for all t. Suppose there exists a k such that $\tilde{l}_k^{\gamma} < \frac{1}{1+\gamma}$. Let total revenues $Z(l) \equiv l - l^{1+\gamma}$. The maximum value of $Z(\cdot)$ is achieved at $l^* = \frac{1}{(1+\gamma)^{\frac{1}{\gamma}}}$. Since $Z(\infty) = -\infty < Z(\tilde{l}_k) < Z(l^*)$, applying the Intermediate Value Theorem, we can find a $l_k^{'} > \frac{1}{(1+\gamma)^{\frac{1}{\gamma}}} > \tilde{l}_k$ such that $Z(l_k^{'}) = Z(\tilde{l}_k)$. Construct an alternative sequence of $\left\{\tilde{B}_t, \hat{l}_t\right\}_t$ where $\hat{l}_t = l_k^{'}$ for $t = k$ and $\hat{l}_t = \tilde{l}_t$ for $t \neq k$. This sequence $\left\{\tilde{B}_t, \hat{l}_t\right\}_t$ also satisfies constraints (9) for all t but has a strictly higher welfare as the objective function (10) is strictly increasing in l_t. Thus, we obtain a contradiction that $\{\tilde{B}_t, \tilde{l}_t\}$ is a optimal solution.

Since $\tilde{l}_t^{\gamma} \geq \frac{1}{1+\gamma}$, it implies that $\tilde{Z}_t \leq \gamma \left(\frac{1}{1+\gamma}\right)^{1+1/\gamma}$, which is our definition of $\overline{Z}$, and the relationship between $\tilde{Z}_t$ and $\tilde{l}_t$ and $\tilde{\tau}_t$ and $\tilde{Z}_t$ is one to one in the relevant range of $\tilde{Z}_t$. Since the Ψ satisfies $\Psi(l - l^{1+\gamma}) = \frac{1}{1+\gamma} l^{1+\gamma}$ and defined for $l^{\gamma} \geq \frac{1}{1+\gamma}$. Differentiate it twice and take the limits as $l^{\gamma} \to \frac{1}{1+\gamma}$ and $l^{\gamma} \to \infty$ (which corresponds to $Z \to \bar{Z}$ and $Z \to -\infty$) to show properties of Ψ. □

For our analysis we need the optimality conditions to (12):

$$V'(B\_) = \beta \int R(s) V'\left(\tilde{B}(s, B\_)\right) \pi(ds) - \bar{\kappa} + \underline{\kappa} \tag{53}$$

and

$$\gamma \Psi'\left(\tilde{Z}(s, B\_)\right) = -1 + \beta V'\left(\tilde{B}(s, B\_)\right) - \bar{\kappa}(s) + \underline{\kappa}(s), \tag{54}$$

where $\bar{\kappa}(s)$ and $\underline{\kappa}(s)$ are the Lagrange multipliers on $B \leq \overline{B}$ and $B \geq \underline{B}$, and $\bar{\kappa} = \int \bar{\kappa}(s) R(s) \pi(ds)$, $\underline{\kappa} = \int \underline{\kappa}(s) R(s) \pi(ds)$.

Proof of Lemma 2

Proof. The right side of expression (12) can be maximized separately for each s. As the objective function being additively separable and concave in (Z, B) and the constraint (13) takes the form $Z + B = E(s, B\_)$, we can apply Corollary 2(ii) in Quah (2007) to conclude that optimal values of $Z(s), B(s)$ are increasing in $E(s, B\_)$. We next show that the solution is strictly increasing when it is interior. Suppose that $E(s'', B\_) > E(s', B\_)$ but $\tilde{B}(s'', B\_) = \tilde{B}(s', B\_)$. Then $\tilde{Z}(s'', B\_) > \tilde{Z}(s', B\_)$ from (13). Strict concavity of $\Psi(\cdot)$ and an interior solution to (54) then implies

$$-1 + \beta V'(\tilde{B}(s', B\_)) = -1 + \beta V'(\tilde{B}(s'', B\_)) = \gamma \Psi'(\tilde{Z}(s'', B\_)) < \gamma \Psi'(\tilde{Z}(s', B\_)),$$

contradicting (54) when solution is interior. Therefore $\tilde{B}(s'', B\_) > \tilde{B}(s', B\_)$. Strict monotonicty of $\tilde{Z}$ and $\tilde{\tau}$ then follows from (54) and the fact that $\tilde{\tau}$ is an increasing function of $\tilde{Z}$. $\square$

Proof of Proposition 1

For the proof of Proposition 1 we will use the observation:

Lemma 6. *The random variables $p(s)$ and $g(s)$ are perfectly correlated if and only if there exists a B such that $E(s, B\_)$ is independent of s.*

Proof. If p and g are perfectly correlated we can write them as $p(s) = \bar{p} + \psi g(s)$ for some $\psi \neq 0$. Set $B = -\frac{\beta\bar{p}}{\psi}$ and observe that $E(s, B) = -\frac{\bar{p}}{\psi}$. The "only if" part follows immediately from the definition of $E(s, B)$. $\square$

To show existence of the invariant distribution we consider two cases that depend on whether or not $p(s)$ and $g(s)$ are perfectly correlated. In cases where they are, the arguments in the proof of Proposition 2 show $\lim_{t\to\infty} \tilde{B}_t = B^*$ and thus the (unique) invariant distribution exists with all mass on B^*.

Now we deal with the case $p(s)$ and $g(s)$ are not perfectly correlated. Our strategy is use to use Theorem 2 in Hopenhayn and Prescott (1992) and a crucial step is to establish that the probability of reaching $\left[\overline{B} - \epsilon, \overline{B}\right]$ starting from any $B\_ \in \left[\underline{B}, \overline{B} - \epsilon\right]$ in a finite number of steps is bounded from below. To show that we will use the next lemma.

Lemma 7. *There are sets $S', S'' \subset S$ of positive measure such that $\tilde{B}(s', B\_) \geq B\_$ for all $s' \in S'$ and $B\_ \geq \tilde{B}(s'', B\_)$ for all $s'' \in S''$ with at least one inequality strict and both inequalities strict if $B\_ \in \left(\underline{B}, \overline{B}\right)$.*

Proof. Suppose $\overline{B} > B\_ \geq \tilde{B}(s, B\_)$ for almost all s, which by strict concavity of V implies

$$V'(B\_) \quad \leq \quad \beta \mathbb{E} R(\cdot) V'\left(\tilde{B}(\cdot, B\_)\right). \tag{55}$$

From Lemma 2, except when $B\_ = \tilde{B}(s, B\_) = \underline{B}$ for almost all s, the inequality (55) can hold with an equality only if $E(s, B\_)$ is independent of s. This is ruled out when $p(s)$ and $g(s)$ are not perfectly correlated and thus $V'(B\_) < \beta \mathbb{E} R(\cdot) V'\left(\tilde{B}(\cdot, B\_)\right)$. On the other hand equation (53) implies that $V'(B\_) \geq \beta \mathbb{E} R(\cdot) V'\left(\tilde{B}(\cdot, B\_)\right)$, a contradiction. Thus, if $\overline{B} > B\_$ then there exists $S' \subset S$ with positive measure such that $\tilde{B}(s', B\_) > B\_$ for all $s' \in S'$. Analogous arguments show that if $B\_ > \underline{B}$ then there exits S'' with positive measure such that $\tilde{B}(s'', B\_) < B\_$ for all $s'' \in S''$ and that there exist s' and s'' such that $B(s', \overline{B}) = \overline{B}$ and $\tilde{B}(s'', \underline{B}) = \underline{B}$. $\square$

Lemma 7 implies that both $\mathbb{E}\left[\tilde{B}(s, B\_) - B\_ \middle| \tilde{B}(s, B\_) \geq B\_\right]$ and $\Pr\left\{\tilde{B}(s, B\_) \geq B\_\right\}$ are positive for all $B\_ \in [\underline{B}, \overline{B} - \epsilon]$ for any $\epsilon > 0$. As both of these terms are continuous functions of $B\_$, compactness of $[\underline{B}, \overline{B} - \epsilon]$ implies that there exists $\underline{d}, \underline{p} > 0$ such that

$$\mathbb{E}\left[\tilde{B}(s, B\_) - B\_ \middle| \tilde{B}(s, B\_) \geq B\_\right] \geq \underline{d} \tag{56}$$

and

$$\Pr\left\{\tilde{B}(s,B\_)\geq B\_\right\}\geq \underline{p} \tag{57}$$

for all $B\_\in[\underline{B},\overline{B}-\epsilon]$. As $\tilde{B}(s,B\_)-B\_$ is bounded above by $\overline{D}=\overline{B}-\underline{B}$, we obtain that[24]

$$\Pr\left\{\tilde{B}(s,B\_)-B\_>\underline{d}/2\right\}\geq\frac{\underline{d}/2}{\overline{D}}\underline{p}$$

for all $B\_\in[\underline{B},\overline{B}-\epsilon]$. Therefore, there must exist an integer n and $\varrho>0$ that that probability of reaching $\left[\overline{B}-\epsilon,\overline{B}\right]$ starting from any $B\_\in\left[\underline{B},\overline{B}-\epsilon\right]$ in n steps is greater than ϱ. Analogous arguments establish that probability of reaching $\left[\underline{B},\overline{B}-\epsilon\right]$ starting from $B\_\in\left[\overline{B}-\epsilon,\overline{B}\right]$ in finite number of steps is finite. Since by Lemma 2 policy functions are monotone in $B\_$, Theorem 2 in Hopenhayn and Prescott (1992) establishes the existence of an unique invariant distribution.

Proof of Proposition 3

Since $R(s)=\frac{p(s)}{\beta\bar{p}}$, we can assume that $R(s)$ is an exogenous stochastic process given by $R(s)=\frac{1}{\beta}\left[1+\sigma\epsilon_R(s)\right]$ for some mean zero random variable ϵ_R. Let $\boldsymbol{\epsilon}=(\epsilon_g,\epsilon_R)$ and write policy functions as $\tilde{B}(\sigma\boldsymbol{\epsilon},B\_;\sigma)$ and $\tilde{Z}(\sigma\boldsymbol{\epsilon},B\_;\sigma)$. Let $\tilde{\mu}(B\_;\sigma)\equiv V'(B\_;\sigma)$. When the solution is interior, the first order condition with respect to Z, equation (54) implies that

$$\tilde{Z}\left(\sigma\boldsymbol{\epsilon},B\_;\sigma\right)=\Psi'^{-1}\left(\frac{-1+\beta\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon},B\_;\sigma\right);\sigma\right)}{\gamma}\right)\equiv\Phi\left(\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon},B\_;\sigma\right);\sigma\right)\right). \tag{58}$$

Therefore we can write equations (13) and (53) as

$$\frac{B\_}{\beta}\left(1+\sigma\epsilon_R\right)+\bar{g}+\sigma\epsilon_g=\Phi\left(\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon},B\_;\sigma\right);\sigma\right)\right)+\tilde{B}\left(\sigma\boldsymbol{\epsilon},B\_;\sigma\right)\ \text{ for all }\boldsymbol{\epsilon},B\_,\sigma \tag{59}$$

and

$$\tilde{\mu}(B\_;\sigma)=\mathbb{E}\left[\left(1+\sigma\epsilon_R\right)\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon},B\_;\sigma\right);\sigma\right)\right]\ \text{ for all B\_,}\sigma. \tag{60}$$

We use $\tilde{B}_g,\tilde{B}_R,\tilde{B}_{B\_},\tilde{B}_\sigma$ to denote derivatives of $\tilde{B}(\sigma\boldsymbol{\epsilon},B\_;\sigma)$ with respect to its first, second, third and fourth argument evaluated at $(\mathbf{0},B\_;0)$. All the second derivatives and derivatives of $\tilde{\mu}$ evaluated at $(\mathbf{0},B\_;0)$ are defined analogously. Finally we use $\bar{\Phi}$ and Φ_μ to denote the value and derivative of $\Phi(\cdot)$ evaluated at $\tilde{\mu}\left(B\_;0\right)$. Observe that the optimality requires that $\tilde{B}\left(\mathbf{0},B\_;0\right)=B\_$.

To calculate these derivatives, differentiate (59) and (60) with respect to $\epsilon_g,\epsilon_R,B\_$ and

[24] An arbitrary random variable $f(s,B\_)$, standing in for the $\tilde{B}(s,B\_)\geq B\_$, that minimizes probability of $\Pr\left\{\tilde{B}(s,B\_)-B>\frac{d}{2}\right\}$ while still satisfying equations (56) and (57) is obtained by placing a mass ψ on $\overline{B}-B\_$ and a mass $\Pr\left\{\tilde{B}(s,B\_)-B\_\geq 0\right\}-\psi$ on $\frac{d}{2}$. As $\overline{B}-B\_\leq\overline{D}$, ψ can be bounded from below by $\frac{d/2}{\overline{D}-\underline{d}/2}\underline{p}$.

evaluate them as $\sigma \to 0$ to obtain

$$\frac{B\_}{\beta} = \left[\Phi_\mu \tilde{\mu}_{B\_} + 1\right] \tilde{B}_R, \quad 1 = \left[\Phi_\mu \tilde{\mu}_{B\_} + 1\right] \tilde{B}_g,$$
$$\frac{1}{\beta} = \left[\Phi_\mu \tilde{\mu}_{B\_} + 1\right] \tilde{B}_{B\_}, \quad \tilde{\mu}_{B\_} = \tilde{\mu}_{B\_} \tilde{B}_{B\_}.$$

These equations can be solved for

$$\tilde{B}_{B\_} = 1, \ \Phi_\mu \tilde{\mu}_{B\_} + 1 = \frac{1}{\beta}, \ \tilde{B}_R = B\_, \ \tilde{B}_g = \beta. \tag{61}$$

Using the same steps for the second order derivatives we can show that

$$\begin{gathered} \tilde{B}_{B\_B\_} = \tilde{B}_{gg} = \tilde{B}_{RR} = \tilde{B}_{Rg} = 0, \\ \Phi_{\mu\mu} \tilde{\mu}^2_{B\_} + \Phi_\mu \tilde{\mu}_{B\_B\_} = 0. \end{gathered} \tag{62}$$

Similarly, differentiating (59) and (60) with respect to σ and its cross-partials we get

$$\begin{aligned} &\tilde{\mu}_\sigma = \tilde{B}_\sigma = \tilde{B}_{\sigma p} = \tilde{B}_{\sigma g} = 0, \\ &\beta \mu_{\sigma\sigma} \Phi_\mu + B_{\sigma\sigma} = 0, \\ &\tilde{B}_{\sigma\sigma} = -2\mathbb{E}\left[\epsilon_R \left(\epsilon_R B\_ + \beta \epsilon_g\right) + \frac{\bar{\Phi}(1-\beta)}{\beta \left(\Phi_\mu\right)^2} \left(\epsilon_R B\_ + \beta \epsilon_g\right)^2\right]. \end{aligned} \tag{63}$$

Using these expressions and after dropping terms that are zero, the second order Taylor expansion of $\tilde{B}\left(\sigma \boldsymbol{\epsilon}, B\_; \sigma\right)$ gives

$$\begin{aligned} \tilde{B}\left(\sigma \boldsymbol{\epsilon}, B\_; \sigma\right) &= B\_ + \tilde{B}_g \sigma \epsilon_g + \tilde{B}_R \sigma \epsilon_R + \frac{\sigma^2}{2} \tilde{B}_{\sigma\sigma} + \mathcal{O}(\sigma^3) \\ &= B\_ + \beta \sigma \epsilon_g + B\_ \sigma \epsilon_R - \operatorname{var}\left(\epsilon_R\right) \sigma^2 B\_ - \beta \operatorname{cov}(\epsilon_R, \epsilon_g) \sigma^2 + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2). \end{aligned} \tag{64}$$

Substituting $\sigma \epsilon_g = g(s) - \bar{g}$, $\sigma \epsilon_R = R(s) - \beta^{-1}$ into this expression we obtain (17). Use the definition of B^* and the fact that $1 - x\sigma^2 = \frac{1}{1+x\sigma^2} + \mathcal{O}(\sigma^3)$ and $\sigma = \frac{\sigma}{1+\sigma^2 x} + \mathcal{O}(\sigma^3)$ for any constant x we can rewrite (64) as

$$\begin{aligned} \tilde{B}\left(s, B\_\right) = B^* + \beta\left[g(s) - \bar{g}\right] + B^*\left[\beta R(s) - 1\right] + \frac{\beta R(s) - 1}{1 + \beta^2 \operatorname{var}(R)} (B\_ - B^*) \\ + \frac{(B\_ - B^*)}{1 + \beta^2 \operatorname{var}(R)} + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2), \end{aligned}$$

which immediately gives the mean and the speed of mean-reversion of $\tilde{B}$ in Proposition 3. To compute ergodic variance, use the law of total variance.

$$
\begin{aligned}
\mathrm{var}(\tilde{B}_t) &= \mathbb{E}\left(\mathrm{var}\left[\tilde{B}(\sigma\boldsymbol{\epsilon}_t, B_{t-1}) - B^*|B_{t-1}\right]\right) + \mathrm{var}\left(\mathbb{E}\left[\tilde{B}(\sigma\boldsymbol{\epsilon}_t, B_{t-1}) - B^*|B_{t-1}\right]\right) \\
&= \mathbb{E}\left(\beta^2\mathrm{var}(RB^* + g) + \frac{\beta^2\mathrm{var}(R)(B_{t-1} - B^*)^2}{(1+\beta^2\mathrm{var}(R))^2}\right) + \mathrm{var}\left(B^* + \frac{(B_{t-1} - B^*)}{(1+\beta^2\mathrm{var}(R))}\right) + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2) \\
&= \beta^2\mathrm{var}(RB^* + g) + \frac{\beta^2\mathrm{var}(R)\mathrm{var}(\tilde{B}_{t-1})}{(1+\beta^2\mathrm{var}(R))^2} + \frac{\mathrm{var}(\tilde{B}_{t-1})}{(1+\beta^2\mathrm{var}(R))^2} + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2).
\end{aligned}
$$

By definition in the invariant distribution $\mathrm{var}(\tilde{B}_t) = \mathrm{var}(\tilde{B}_{t-1})$ and replacing $\frac{1}{1+\beta^2\mathrm{var}(R)} = 1-\beta^2\mathrm{var}(R)+\mathcal{O}(\sigma^3,(1-\beta)\sigma^2)$ we get the expression for ergodic variance of $\tilde{B}_t$ in Proposition 3.

For $\tilde{Z}_t$ we exploit the fact that $\tilde{Z}(s, B\_) = \Phi\left(\mu\left(s, B\_\right)\right)$ and obtain the second order Taylor expansion of $\Phi\left(\mu\left(s, B\_\right)\right)$ as

$$
\begin{aligned}
&\Phi\left(\mu\left(s, B\_\right)\right) = \Phi^* + \left(\frac{1-\beta}{\beta}\right)\left(B\_ - B^*\right) + \left(\frac{1-\beta}{\beta}\right)\beta\left[g(s) - \bar{g}\right] + \left(\frac{1-\beta}{\beta}\right)B^*\left[\beta R(s) - 1\right] + \\
&\left(\frac{1-\beta}{\beta}\right)\left(B\_ - B^*\right)\left[\beta R(s) - 1\right] + \beta\left(B\_ - B^*\right)\mathrm{var}\left(R\right) + \mathcal{O}\left(\sigma^3, (1-\beta)\sigma^2\right),
\end{aligned}
$$

where $\Phi^* = \left(\frac{1-\beta}{\beta}\right)B^* + \bar{g}$. The steps to compute the ergodic moments of $\tilde{Z}_t$ that are reported in Proposition 3 are similar to those of $\tilde{B}_t$ and hence we omit them.

A.2 Proof of Proposition 4

Using strict concavity of U and Ψ, standard arguments show that value function characterizing the optimal plan with transfers in equation (19) is continuous, strictly concave, and differentiable on $\left[\underline{B}, \overline{B}\right]$. The steps to show uniqueness of invariant distribution of $\left\{\tilde{B}_t, \tilde{Z}_t, \tilde{T}_t\right\}$ are identical to those in the proof of Proposition 1 as in Appendix A.1. They follow from the concavity of the value function and the properties that optimal policy rules $\tilde{B}\left(\cdot,\cdot\right)$, $\tilde{Z}\left(\cdot,\cdot\right)$ and $-\tilde{T}(\cdot,\cdot)$ are continuous and increasing in $B\_$ for all s and strictly increasing in E in the interior as in Lemma 2.

The first order conditions with respect to $Z(s), T(s), B(s)$ are

$$
\gamma\Psi'\left(Z(s)\right) = \mu(s), \tag{65}
$$

$$
\frac{\omega}{1+n}U_c\left(\frac{T(s)}{1+n}\right) + \frac{1}{1+n} + \mu(s) = 0, \tag{66}
$$

$$
-1 + \beta V'\left(B(s)\right) - \bar{\kappa}\left(s\right) + \underline{\kappa}\left(s\right) = \mu(s), \tag{67}
$$

where $\mu(s)$ is the multiplier on constraint (20) and $\bar{\kappa}\left(s\right), \underline{\kappa}\left(s\right)$ are the multipliers on the boundaries $\left[\underline{B}, \overline{B}\right]$. Equations (65) and (66) allow us to solve for net-tax revenues, $I(s) \equiv Z(s) - T(s)$ as a function of $\mu(s)$

$$
I(\mu) = \Psi'^{-1}\left(\frac{\mu}{\gamma}\right) - (1+n)U_c^{-1}\left(-\frac{1}{\omega} - \frac{1+n}{\omega}\mu\right). \tag{68}
$$

The equilibrium conditions for an interior solution to the planners problem can then be written as

$$\frac{B\_p(s)}{\beta\bar{p}} + g(s) = I\left(\tilde{\mu}\left(\tilde{B}(s, B\_)\right)\right) + \tilde{B}(s, B\_), \tag{69}$$

$$\tilde{\mu}(B\_) = \mathbb{E}\left[\left(\frac{p(s)}{\bar{p}}\right)\tilde{\mu}\left(\tilde{B}(s, B\_)\right)\right]. \tag{70}$$

Note that the system of equations (69) and (70) exactly mirrors the system of equations (59) and (60) with Ψ replaced with I. The approximation of $\tilde{B}_t$ in the proof of Proposition 3 (see Appendix A.1) does not rely on the properties of Ψ, and thus the approximation of the optimal policies $\tilde{B}_t$ and $\tilde{I}_t$ that solve (69) and (70) will be identical to the representative agent case. Importantly, the approximation to the ergodic distribution of $\tilde{B}_t$ and $\tilde{I}_t$ is independent of the value of ω.

Define the level of transfers when the government has a net tax revenue of I as $T(I;\omega)$. Combine the first order conditions, (65) and (66) and substitute $Z = I + T$ to express transfers, $T(I;\omega)$ implicitly using

$$\frac{\omega}{1+n}U_c\left(\frac{T(I;\omega)}{1+n}\right) + \gamma\Psi'(I + T(I;\omega)) = -\frac{1}{1+n}.$$

Differentiating with respect to ω we get,

$$\frac{\partial T}{\partial \omega}(I;\omega) = -\frac{\frac{1}{1+n}U_c\left(\frac{T(I;\omega)}{1+n}\right)}{\frac{\omega}{(1+n)^2}U_{cc}\left(\frac{T(I;\omega)}{1+n}\right) + \gamma\Psi''(I + T(I;\omega))}$$

Both $U_{cc} < 0$ and $\Psi'' < 0$, we can conclude that $\frac{\partial T}{\partial \omega}(I;\omega) > 0$. Thus $T_t = T(I_t;\omega)$ is increasing in ω and first order stochastic dominance follows immediately.

Finally, when we modify the tax scheme to allow transfers only to unproductive agents, the first order conditions are unchanged except equation (66) that is modified to

$$\frac{\omega}{n}U_c\left(\frac{T(s)}{n}\right) + \frac{1}{n} + \mu(s) = 0$$

and all the subsequent arguments remain the same.

A.3 Proofs for Section 3.2

Proof for Lemma 3

Problem (22) maintains the structure of a concave objective function and linear constraints so using the same arguments as in the single asset case we can conclude that $V(B\_)$ is differentiable and strictly concave on the domain $\left[\underline{B}, \overline{B}\right]$ and consequently the optimal policies $\tilde{B}(s, B\_)$ and $\tilde{Z}(s, B\_)$ are unique. The next lemma shows that there is a unique optimal portfolio choice $\tilde{\mathbf{B}}(B\_)$.

Lemma 8. *If the random variables* $\mathbf{R}$ *are linearly independent, then there exists a unique portfolio associated with policy rules* $B(s, B\_)$ *and* $Z(s, B\_)$.

Proof. Suppose that there exists two optimal portfolio rules $\left\{B^{\dagger,k}(B\_)\right\} \neq \left\{B^{*,k}\_(B\_)\right\}$ such that $\sum_k B^{\dagger,k}(B\_) = \sum_k B^{*,k}(B\_) = B\_$. Substituting in budget constraint (23) and taking differences we obtain that for all states $s \in S$

$$\sum_k \left[\left\{B^{\dagger,k}(B\_)\right\} - \left\{B^{*,k}\_(B\_)\right\}\right] R^k(s) = 0.$$

This contradicts the assumption that $\mathbf{R}$ are linearly independent. ◻

The gross government expenditures, $E(s, B\_)$ with multiple assets are

$$E(s, B\_) = \mathbf{R}^T(s)\tilde{\mathbf{B}}(B\_) + g(s).$$

Lemma 2 extends to the policy rules for total assets $\tilde{B}(s, B\_)$ and tax revenues $\tilde{Z}(s, B\_)$. To show uniqueness of the invariant distribution we will split the problem into two cases depending on whether $g(s)$ is in the span of $\mathbf{R}$.

We begin when $g(s)$ is not in the span of $\mathbf{R}$. Then $\text{var}(E(\cdot, B\_)) > 0$ for all $B\_$, and is uniformly bounded from below by a positive constant. We can then immediately apply our results from the single asset case to show that there exists a unique invariant distribution using Theorem 2 in Hopenhayn and Prescott (1992).

When $g(s)$ is in the span of $\mathbf{R}$, the mixing condition required by Hopenhayn and Prescott (1992) fails. Instead, we shall prove the existence of unique invariant distribution along the lines of Proposition 2 by showing that $\tilde{B}_t \to B^*$ almost surely. Without loss of generality assume that

$$R^1(s) = \frac{1}{\beta} - \psi\bar{g} + \psi g(s)$$

for some non zero ψ and that the remaining R^k satisfy the orthogonality condition $\mathbb{E}(R^k|\mathbf{R}^{-j}) = \frac{1}{\beta}$ for all j and k where $\mathbf{R}^{-j}$ refers to the vector of returns with out the j^{th} return.[25] $E(s, B\_)$ can then be written as

$$E(s, B\_) = \sum_{k\geq 2} R^k(s)\tilde{B}^k(B\_) + \tilde{B}^1(B\_)\left(\frac{1}{\beta} - \psi\bar{g}\right) + g(s)(1 + \psi\tilde{B}^1(B\_)). \tag{71}$$

Since $g(s)$ is in the span of $\mathbf{R}$ we set $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ to be non-singular. From equation (21), $B^* = -\frac{1}{\psi}$ and set $\tilde{B}^1(B^*) = B^*$, $\tilde{B}^k(B^*) = 0$ for $k \geq 2$ to get that $\tilde{B}(s, B^*) = B^*$ for all s. Thus B^* is a steady state, it remains to be shown that $\tilde{B}_t \to B^*$ almost surely.

The Euler equation for each of the k security

$$\begin{aligned} V'(B\_) = \beta\mathbb{E}[R^k(\cdot)V'(\tilde{B}(\cdot, B\_))] &= \mathbb{E}[V'(\tilde{B}(\cdot, B\_))] + \text{cov}(R^k(\cdot), V'(\tilde{B}(\cdot, B\_)) \\ &= \mathbb{E}[V'(\tilde{B}(\cdot, B\_))] + \mathbb{E}\left[\text{cov}\left(R^k(\cdot), V'(\tilde{B}(\cdot, B\_)|\mathbf{R}^{-k}\right)\right]. \end{aligned}$$

In the last equality we used the law of total covariance combined with $\mathbb{E}(R^k|R^j) = \frac{1}{\beta}$ for $k, j \neq 1$.

Let $B\_ < B^*$ and suppose that $\tilde{B}^1(B\_) \geq B^*$. Equation (71) tells us that, conditional on $\mathbf{R}^{-1}$, $E(\cdot, B\_)$ is perfectly positively correlated with $R^1(s)$. As $\tilde{B}$ is increasing in E and V' is

[25] In effect we choose an orthogonal basis of the space spanned by the returns.

a decreasing function we can immediately conclude that $\mathbb{E}\left[\operatorname{cov}\left(R^1(\cdot), V'(\tilde{B}(\cdot, B\_)|\mathbf{R}^{-1}\right)\right] \leq 0$. However, $\tilde{B}^1(B\_) \geq B^*$ implies that $\tilde{B}^k < 0$ for some $k \geq 2$. As conditioning on $\mathbf{R}^{-k}$ implies conditioning on g,[26] we can immediately conclude that, conditional on $\mathbf{R}^{-k}$ $E(\cdot, B\_)$ is perfectly negatively correlated with R^k. The logic from above then implies that $\mathbb{E}\left[\operatorname{cov}\left(R^k(\cdot), V'(\tilde{B}(\cdot, B\_)|\mathbf{R}^{-k}\right)\right] > 0$, a contradiction as all K Euler equations must hold. We can therefore conclude that for $B\_ \leq B^*$ that $\tilde{B}\_^1(B\_) \leq B^*$, and hence, $\mathbb{E}\left[\operatorname{cov}\left(R^1(\cdot), V'(\tilde{B}(\cdot, B\_)|\mathbf{R}^{-1}\right)\right] \geq 0$. Applying this to the Euler equation allows us to conclude that if $B_{-1} < B^*$ then $\tilde{B}_t \leq B^*$ for all t and

$$V'(\tilde{B}_t) \geq \mathbb{E}V'(\tilde{B}_{t+1}).$$

Using the same steps as the proof of Proposition 2 we can apply the martingale convergence theorem to conclude that $\tilde{B}_t \to B^*$ almost surely. The case when $B_{-1} > B^*$ is symmetric.

Lastly we show that $g(s)$ in the span of $\mathbf{R}$ and $\mathbb{C}(\mathbf{R}, \mathbf{R})$ singular characterizes an optimal policy where tax revenues are constant and total debt is equal to its initial value for all dates.

Lemma 9. *$g(s)$ is in the span of $\mathbf{R}$ and $\mathbb{C}(\mathbf{R}, \mathbf{R})$ is singular if and only if the optimal rule for total debt satisfies $\tilde{B}(s, B\_) = B\_$ for $B\_ \in [\underline{B}, \overline{B}]$.*

Proof. In the proof we will use the observation that in absence of redundant securities, the matrix $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ is singular, if and only if there exists a linear combination of $\mathbf{R}$ that yields a risk free return. Now if $g(s)$ is in the span of $\mathbf{R}$ then there exist a non-zero vector $\boldsymbol{\psi}_1$ such that $g(s) = \sum_k \psi_1^k R^k(s)$ and with singular $\mathbb{C}[\mathbf{R}, \mathbf{R}]$ there exists a exist a non-zero vector $\boldsymbol{\psi}_2$ such that $\sum_k \psi_2^k R^k(s) = \frac{1}{\beta}$. For any initial assets $B_{-1} \in [\underline{B}, \overline{B}]$, consider the portfolio that allocates

$$B^k = \left(\frac{B\_ + \sum_k \psi_1^k}{\sum_k \psi_2^k}\right) \psi_2^k - \psi_1^k$$

in security k. It is easy to check that $\sum_k B^k = B\_$ and now we verify that $\mathbf{R}^T(s)\mathbf{B} + g(s)$ is independent of s:

$$\mathbf{R}^T(s)\mathbf{B} + g(s) = \sum_k R^k(s)\left[\left(\frac{B\_ + \sum_k \psi_1^k}{\sum_k \psi_2^k}\right)\psi_2^k - \psi_1^k\right] + g(s) = \frac{1}{\beta}\left(\frac{B\_ + \sum_k \psi_1^k}{\sum_k \psi_2^k}\right).$$

Thus the policy rule $\tilde{B}(s, B\_) = B\_$ and $\tilde{Z}(s, B\_) = \left(\frac{1-\beta}{\beta}\right) B\_ + \bar{g}$ is satisfy the budget constraint (23). As our setting with K securities and returns $\mathbf{R}$ is a restricted version of a complete markets problem[27] where policies $\tilde{B}(s, B\_) = B\_$ and $\tilde{Z}(s, B\_) = \left(\frac{1-\beta}{\beta}\right) B\_ + \bar{g}$ are optimal, we can conclude that the same policies are optimal for (22) too.

Now we show the converse. Suppose $\tilde{B}(s, B\_) = B\_$ for all $B\_$. The budget constraint implies that

$$g(s) = B\_ + Z^*(B\_) - \sum_k B^k R^k(s) \tag{72}$$

Now set $B\_ = -\beta\mathbb{E}g$ and equation (72) gives

[26] As R^1 is perfectly correlated with g

[27] By this we mean a market setting with a complete set of Arrow securities as in Lucas and Stokey (1983)

$$\begin{aligned} g(s) &= \sum_k B_k + \frac{1-\beta}{\beta}\sum_k B_k + \mathbb{E}g - \sum_k B^k R^k(s) \\ g(s) &= \sum_k B^k R^k(s) \end{aligned}$$

This shows that $g(s)$ is in the span of $\mathbf{R}$. For $B'\_ \neq B''\_$ and respective optimal portfolios $\left\{B'^k\right\}$, $\left\{B''^k\right\}$ equation (72) again implies that

$$0 = B'\_ - B''\_ + Z^*(B'\_) - Z^*(B''\_) - \sum_k \left(B'^k - B''^k\right) R^k(s) \tag{73}$$

Since $\left(B'^k - B''^k\right)$ cannot be a zero vector, we see from (73) that a linear combination of $\mathbf{R}$ yields a constant. Thus a risk free bond is in the space spanned by the returns $\mathbf{R}$ and $\mathbb{C}[\mathbf{R},\mathbf{R}]$ is singular. □

Proof of Lemma 4 and Proposition 5

Let $\boldsymbol{\epsilon} \equiv (\epsilon_g, \epsilon_{\mathbf{R}})$ and functions $\tilde{B}(\sigma\epsilon, B\_;\sigma)$, $\tilde{\mathbf{B}}(B\_;\sigma)$ and $\tilde{\mu}(B\_;\sigma)$ solve the optimality conditions for Problem (22) with K assets:

$$\beta^{-1}\left[\mathbf{1} + \sigma\epsilon_{\mathbf{R}}\right]^T \tilde{\mathbf{B}}(B\_;\sigma) = \Phi\left(\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon}, B\_;\sigma\right)\right)\right) - \bar{g} - \sigma\epsilon_g + \tilde{B}\left(\sigma\boldsymbol{\epsilon}, B\_;\sigma\right), \tag{74}$$

$$\tilde{\mu}(B\_) = \mathbb{E}\left[\left(1 + \sigma\epsilon_{R^k}\right)\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon}, B\_;\sigma\right)\right)\right] \text{ for all } k, \tag{75}$$

$$\mathbf{1}^T\tilde{\mathbf{B}}(B\_;\sigma) = B\_. \tag{76}$$

The proof will follow the similar steps and notation convention as used in Appendix A.1. With $K > 1$ equation (74)-(76) has multiple solutions for $\tilde{\mathbf{B}}(B\_;0)$ at $\sigma = 0$. However uniqueness of policy rules as asserted in Lemma 3 implies that $\lim_{\sigma\to 0}\tilde{\mathbf{B}}(B\_;\sigma)$ exists, which for now we denote by $\mathbf{B}^k(B\_)$, and then show below that the existence of second derivatives of $\tilde{B}\left(\sigma\boldsymbol{\epsilon}, B\_;\sigma\right)$ along with equations (74)-(76) implies that this limiting portfolio is given by $\mathbf{B}^*(B\_)$ as required by Lemma 4.

Many of the first order and second order terms including the the steps to obtain them are very similar to the single asset case. We get[28]

$$\begin{aligned} &\tilde{B}_{R^k}(B\_) = B^k(B\_) \\ &\tilde{B}_\sigma(B\_) = \tilde{\mu}_\sigma(B\_) = \tilde{B}^k_\sigma(B\_) = 0. \end{aligned}$$

[28]The derivatives $\tilde{B}^k_{B\_}(B\_)$ are undetermined, but this does not affect any future calculations or the approximated policy rules for total assets $\tilde{B}$.

For the second order terms,

$$
\begin{aligned}
&\tilde{B}_{B\_B\_}(B\_) = 0, \tilde{\mu}_{B\_B\_}(B\_) = -\frac{\Phi_{\mu\mu}(B\_)}{\Phi_{\mu}(B\_)}\left(\tilde{\mu}_{B\_B\_}(B\_)\right)^2, \\
&\tilde{B}_{R^kR^k} = \tilde{B}_{gg} = \tilde{B}_{R^kg} = 0 \quad \forall k.
\end{aligned}
$$

The calculations for the second derivative with respect to σ is key to show Lemma 4. Differentiating (75) twice with respect to σ, we obtain

$$
\begin{aligned}
\tilde{\mu}_{\sigma\sigma}(B\_) = \mathbb{E}\Bigg[\tilde{\mu}_{\sigma\sigma}(B\_) + \tilde{\mu}_{B\_}(B\_)\tilde{B}_{\sigma\sigma}(B\_) + 2\epsilon_{R^k}\tilde{\mu}_{B\_}(B\_)\left(\sum_j \tilde{B}_{R^j}(B\_)\epsilon_{R^j} + \tilde{B}_g(B\_)\epsilon_g\right) \\
+ \tilde{\mu}_{B\_B\_}(B\_)\left(\sum_j \tilde{B}_{R^j}(B\_)\epsilon_{R^j} + \tilde{B}_g(B\_)\epsilon_g\right)^2\Bigg].
\end{aligned}
\tag{77}
$$

We can eliminate $\tilde{\mu}_{\sigma\sigma}(B\_)$ from (77), substitute for $\tilde{B}_{R^j}(B\_)$ and solve out for $\tilde{B}_{\sigma\sigma}(B\_)$. For all k we obtain,

$$
\tilde{B}_{\sigma\sigma}(B\_) = -2\mathbb{E}\left[\epsilon_{R^k}\left(\sum_j \epsilon_{R^j}\tilde{B}^j\_(B\_) + \beta\epsilon_g\right) + \frac{\Phi(B\_)(1-\beta)}{\beta\Phi_{\mu}(B\_)}\left(\sum_j \epsilon_{R^j}\tilde{B}^j\_(B\_) + \beta\epsilon_g\right)^2\right].
\tag{78}
$$

System (78) has K equations, one for each security k. To satisfy all of them, there must exist a constant δ and portfolio $\mathbf{B}$ such that

$$
\mathbb{E}\left[\epsilon_{R^k}\left(\sum_j \epsilon_{R^j}\tilde{B}^j\_(B\_) + \beta\epsilon_g\right)\right] = \delta \quad \forall k,
\tag{79}
$$

$$
\mathbf{1}^T\mathbf{B} = B\_.
\tag{80}
$$

Equations (79) corresponds exactly to the first order conditions of the Problem 24 when we set the Lagrange multiplier on the constraint $\mathbf{1}^T\mathbf{B} = B\_$ to be $\frac{2\delta}{\beta^2}$. This shows $\mathbf{B}^k(B\_) = \mathbf{B}^{*,k}(B\_)$ as stated by Lemma 4.

Now we derive an expression for $\mathbf{B}^*(B\_)$ in terms of the primitives and $B\_$. Write equations (79) and (80) as a linear system of equations of the following form

$$
\begin{bmatrix} \beta^2\mathbb{C}(\mathbf{R},\mathbf{R}) & 1 \\ \mathbf{1}^\intercal & 0 \end{bmatrix}\begin{bmatrix} \mathbf{B}^* \\ -\delta \end{bmatrix} = \begin{bmatrix} -\beta^2\mathbb{C}(\mathbf{R},g) \\ B\_ \end{bmatrix}
\tag{81}
$$

where $\mathbf{1}$ is a K dimensional vector of ones. There are two possible types of solutions to (81). If $\mathbb{C}(\mathbf{R},\mathbf{R})$ is not of full rank, the minimization problem (24) has multiple solutions and further the minimum is independent of $B\_$. In these cases $\delta(B\_) = 0$ for all $B\_$. The other case, when $\mathbb{C}(\mathbf{R},\mathbf{R})$ is invertible, we can express $\delta(B\_)$ and $\mathbf{B}^*(B\_)$ as functions of $B\_$. Define a scalar

$\eta \equiv \beta^{-\mathbf{2}}\mathbf{1}^{\intercal}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}\mathbf{1}$, and using equation (81) we have

$$\delta(B\_) = \frac{1}{\eta}\left(B\_ + \mathbf{1}^{\intercal}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}\mathbb{C}(\mathbf{R},g)\right), \tag{82}$$

$$\mathbf{B}^{*}(B\_) = -\left(\beta^{-2}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1} - \frac{\beta^{-4}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}\mathbf{1}\mathbf{1}^{\intercal}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}}{\eta}\right)\beta^{2}\mathbb{C}(\mathbf{R},g) + \frac{\beta^{-2}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}\mathbf{1}B\_}{\eta},$$

which after simplification gives us formula (26) in the text.

After substituting for $\tilde{B}_g(B\_)$, $\left\{\tilde{B}_{R^k}(B\_)\right\}$ and $\tilde{B}_{\sigma\sigma}(B\_)$ in the second order Taylor expansion we obtain

$$\tilde{B}(s,B\_;\sigma) = B\_ + \beta\left[g(s)-\bar{g}\right] + \sum_k B^k\_(B\_)\left[R^k(s)-\beta^{-1}\right] - \delta(B\_) + \mathcal{O}(\sigma^3,(1-\beta)\sigma^2),$$

which after substituting for $\delta(B\_)$ yields (25). Taking unconditional means of equation (25), we see that B^* satisfies $\delta(B^*)=0$ and from equation (82) we can verify the expression for the mean of $\tilde{B}_t$ asserted in Proposition 5. Next taking the mean conditional on $\tilde{B}_t$ on both sides of equation (25) we can verify that the speed of mean reversion equals

$$1-\frac{1}{\eta} = \frac{1}{1+\eta^{-1}} + \mathcal{O}(\sigma^3) = \frac{\beta^{-\mathbf{2}}\mathbf{1}^{\intercal}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}\mathbf{1}}{1+\beta^{-\mathbf{2}}\mathbf{1}^{\intercal}\mathbb{C}(\mathbf{R},\mathbf{R})^{-1}\mathbf{1}}.$$

The steps to compute the ergodic variance of $\tilde{B}$ and the ergodic moments of $\tilde{Z}$ are exactly same as in the proof of Proposition 3 and are omitted.

A.4 Proof of Proposition 6

The proof proceeds along the lines in Proposition 3 and for brevity we will focus on the steps that different. The PV operator defined in the text satisfies the recursion

$$PV(x;s) = x(s) + \beta\mathrm{E}\left(PV(x;s');s\right) \quad \forall s\in S. \tag{83}$$

Without loss of generality, we parametrize the exogenous shock process by defining three mean zero (under the ergodic measure induced by π) random variables $\boldsymbol{\epsilon} = (\epsilon_g, \epsilon_\Theta, \epsilon_R)$ as

$$g(s) \;=\; \bar{g} + \sigma\epsilon_g(s), \quad R(s\_,s) = \frac{1}{\beta}\left(1+\sigma\epsilon_R(s,s\_)\right), \quad \Theta(s) = \bar{\Theta}\left(1+\sigma\epsilon_\Theta(s)\right),$$

where the definition of returns implies that additionally $\mathrm{E}\left(\epsilon_R;s\_\right)=0$ for all $s\_$. The policy rule for debt $\tilde{B}(\sigma\boldsymbol{\epsilon},B\_,s\_;\sigma)$ and $\mu(B\_,s\_;\sigma) = V'(B\_;\sigma)$ satisfy

$$\frac{B\_}{\beta}\left(1+\sigma\epsilon_R\right) + \bar{g} + \sigma\epsilon_g = \bar{\Theta}\left(1+\sigma\epsilon_\Theta\right)\Phi(\tilde{\mu}(\tilde{B}(\sigma\boldsymbol{\epsilon},B\_,s\_;\sigma);\sigma)) + \tilde{B}(\sigma\boldsymbol{\epsilon},B\_,s\_;\sigma), \tag{84}$$

$$\tilde{\mu}(B\_,s\_;\sigma) = \mathrm{E}\left[\left(1+\sigma\epsilon_R\right)\tilde{\mu}\left(\tilde{B}\left(\sigma\boldsymbol{\epsilon},B\_,s\_;\sigma\right);\sigma\right);s\_\right]. \tag{85}$$

We use the same convention as in Appendix A.1 for denoting the derivatives which are now evaluated at $(\mathbf{0}, B\_, s\_; 0)$ and

$$\bar{\Phi} \equiv \Phi(\tilde{\mu}(B\_, s\_; 0)) = \frac{1}{\bar{\Theta}} \left(\frac{1-\beta}{\beta} B\_ + \bar{g} \right).$$

First order terms

Differentiating equation (84) and (85) with respect to $B\_$, $\boldsymbol{\epsilon}$ we obtain

$$\begin{aligned} &\tilde{B}_{B\_} = 1, \quad \tilde{\mu}_{B\_} = \frac{1-\beta}{\beta \Phi_\mu \bar{\Theta}}, \\ &\tilde{B}_R = B\_, \quad \tilde{B}_g = \beta, \text{ and } \tilde{B}_\Theta = -\beta \bar{\Theta} \bar{\Phi}. \end{aligned} \tag{86}$$

We show the calculation of the derivatives with respect to σ as they are different with Markov shocks. First differentiate (84) and (85) with respect to σ to obtain[29]

$$0 = \bar{\Theta} \Phi_\mu (\tilde{\mu}_\sigma + \tilde{\mu}_{B\_} \tilde{B}_\sigma) + \tilde{B}_\sigma = \bar{\Theta} \Phi_\mu \tilde{\mu}_\sigma + \frac{1}{\beta} \tilde{B}_\sigma \tag{87}$$

$$\tilde{\mu}_\sigma = \mathrm{E}\left[\tilde{\mu}_\sigma + \epsilon_R \bar{\mu} + \tilde{\mu}_{B\_} \left(\tilde{B}_\sigma + \tilde{B}_R \epsilon_R + \tilde{B}_g \epsilon_g + \tilde{B}_\Theta \epsilon_\Theta \right); s\_ \right]. \tag{88}$$

Equation (87) implies $\tilde{\mu}_\sigma = -\frac{\tilde{\mu}_B}{1-\beta} \tilde{B}_\sigma$ and using $\mathrm{E}\left(\epsilon_R; s\_\right) = 0$ and $\tilde{B}_R, \tilde{B}_g$ and $\tilde{B}_\Theta$ from equation (86) we get

$$\tilde{B}_\sigma = \mathrm{E}\left[\left(\beta \tilde{B}_\sigma - (1-\beta)\beta \epsilon_g + (1-\beta)\beta \bar{\Theta} \bar{\Phi} \epsilon_\Theta \right); s\_ \right]$$

Now apply the recursive characterization of PV operator from (83) to obtain

$$\tilde{B}_\sigma = -(1-\beta)\beta \mathrm{E}\left[PV(\epsilon_g) - \bar{\Theta} \bar{\Phi} PV(\epsilon_\Theta); s\_ \right]. \tag{89}$$

Second order terms

Differentiating equation (84) with respect to $B\_$ and $\boldsymbol{\epsilon}$ we can show that

$$\begin{aligned} &\tilde{\mu}_{B\_B\_} = -\frac{\Phi_{\mu\mu} (\tilde{\mu}_{B\_})^2}{\Phi_\mu}, \\ &\tilde{B}_{B\_B\_} = \tilde{B}_{gg} = \tilde{B}_{Rg} = \tilde{B}_{RR} = \tilde{B}_{B\_g} = 0 \text{ and } \tilde{B}_{B\_R} = 1. \end{aligned}$$

The cross-partial with respect to TFP shock $\tilde{B}_{\Theta\Theta}, \tilde{B}_{\Theta p}, \tilde{B}_{\Theta g}$, and $\tilde{B}_{\Theta B\_}$ end up contributing $\mathcal{O}((1-\beta)\sigma^2)$ in the Taylor expansion. As in the i.i.d case, except $\tilde{B}_{\sigma\sigma}$, all the cross-partials with respect to σ are zero. Differentiating equation (84) and (85) twice with respect to σ we have

$$\tilde{B}_{\sigma\sigma} = -2(1-\beta) \left[B\_ PV\left(\mathcal{C}^{\epsilon_R, \epsilon_R} \right) + \beta PV\left(\mathcal{C}^{\epsilon_R, PV_{\epsilon_g}} \right) - \beta \bar{\Theta} \bar{\Phi} PV\left(\mathcal{C}^{\epsilon_R, PV_{\epsilon_\Theta}} \right) \right] + \mathcal{O}(1-\beta) \tag{90}$$

Taylor expansion

Combining all the first order and second order terms, we obtain the the approximation to $\tilde{B}$ as in equation (34). Let $\mathrm{var}(x)$ and $\mathrm{cov}(x)$ denote the variance and covariance

[29] The terms associated with ϵ_R, ϵ_Θ and ϵ_g drop out after we substitute for $\tilde{B}_R, \tilde{B}_g$ and $\tilde{B}_\Theta$ from equation (86).

with respect to the ergodic measure induced by π. We will now show that terms $(1-\beta)\left[B\_PV\left(\mathcal{C}^{R,R};s\right)+PV\left(\mathcal{C}^{R,PV(g)};s\right)-\bar{\Phi}PV\left(\mathcal{C}^{R,PV(\Theta)};s\right)\right]$ can be expressed as

$$B\_\mathrm{var}(R)+\mathrm{cov}\left(R,PV(g)\right)-\bar{\Phi}\mathrm{cov}\left(R,PV(\Theta)\right)+\mathcal{O}(\sigma^3,(1-\beta)\sigma^2). \tag{91}$$

Consider the first term $(1-\beta)B\_PV\left(\mathcal{C}^{R,R};s\right)$. Its easy to see that we can express

$$(1-\beta)B\_PV\left(\mathcal{C}^{R,R};s\right)=(1-\beta)B\_PV\left(\mathcal{C}^{R,R}-\mathrm{var}(R);s\right)+B\_\mathrm{var}(R).$$

Our assumptions on π imply that the (linear) operator E has a spectral representation[30], i.e., there exists a sequence of scalars (eigenvalues) $\{\lambda_j\}$ with $1=\lambda_0>\lambda_1>\ldots>\lambda_j>\lambda_{j+1}>\ldots$, and sequence of orthonormal[31] family of random variables (eigenfunctions) $\{\psi_j\}$ such that

$$\mathrm{E}(x;s)=\sum_{j=0}^{\infty}\lambda_j<\psi_j,x>\psi_j(s), \tag{92}$$

where $<x,y>\equiv\mathbb{E}xy$ is the inner product of random variables x and y with respect to the ergodic measure induced by π. It is straightforward to verify that eigenfunction associated with the unit eigenvalue i.e., $\psi_0(s)=1$ and this allows us to express (92) as

$$\mathrm{E}(x;s)=\mathbb{E}x+\sum_{j=1}^{\infty}\lambda_j<\psi_j,x>\psi_j(s).$$

Note that the eigenvalues of the h period ahead conditional expectation operator, $\mathrm{E}^h(x;s)\equiv\mathrm{E}(x(s_{t+h});s)$ are $\left\{\lambda_j^h\right\}$ and the eigenfunctions $\{\psi_j\}$ remain the same as those of E. So for any random variable x,

$$PV(x;s)\quad=\quad\sum_t\beta^t\mathrm{E}^t(x;s)=\frac{\mathbb{E}x}{1-\beta}+\sum_t\beta^t\sum_{j=1}^{\infty}\lambda_j^t<\psi_j,x>\psi_j(s).$$

Thus $(1-\beta)PV(x-\mathbb{E}x;s)=(1-\beta)\sum_t\beta^t\sum_{j=1}^{\infty}\lambda_j^t<\psi_j,x>\psi_j(s)$ and it follows that

$$\|(1-\beta)PV(x-\mathbb{E}x)\|\leq(1-\beta)\sum_t(\beta\lambda_1)^t\,\|\sum_{j=1}^{\infty}<\psi_j,(x-\bar{x})>\psi_j\|=\frac{(1-\beta)}{1-\beta\lambda_1}\|\,(x-\bar{x})\,\|\in\mathcal{O}(1-\beta).$$

For the last equality we use the fact $\lambda_1<1$. The same arguments can be applied to the other terms in equation (34) to obtain equation (35).

To compute the ergodic moments of $\tilde{B}$, define $\mathbb{B}:S\to R$ as

$$\mathbb{B}(s)\equiv B^*-\beta\left[\mathrm{E}\left(PV(g-\bar{g};s');s\right)-\bar{\Phi}(B^*)\mathrm{E}\left(PV(\Theta-\bar{\Theta};s');s\right)\right]$$

and adjusting terms that are of order $\mathcal{O}(\sigma^3)$ using $1-x\sigma^2=\frac{1}{1+x\sigma^2}+\mathcal{O}(\sigma^3)$ and $\sigma=\frac{\sigma}{1+\sigma^2x}+\mathcal{O}(\sigma^3)$

[30]The argument is standard and follows from the fact that E is both compact and self-adjoint. See Dunford and Schwartz (1963, 1966) for more details.

[31]This means $<\psi_j,\psi_j>=1$ and $<\psi_j,\psi_k>=0$ for all $0\leq k<j$.

for any constant x equation (35) can be expressed as

$$\tilde{B}(s,B\_,s\_;\sigma)=\mathbb{B}(s)+\beta\left[PV(g;s)-\mathrm{E}\left(PV(g;s);s\_\right)\right]-\beta\bar{\Phi}\left[PV(\Theta;s)-\mathrm{E}\left(PV(\Theta;s);s\_\right)\right]$$
$$+\mathbb{B}(s)[\beta R(s,s\_)-1]+\left(\frac{1}{1+\beta^2\mathrm{var}(R)}\right)(B\_-\mathbb{B}(s\_))\left[\beta R(s,s\_)-1\right]+\left(\frac{1}{1+\beta^2\mathrm{var}(R)}\right)(B\_-\mathbb{B}(s\_))$$
$$+\mathcal{O}(\sigma^3,(1-\beta)\sigma^2)$$

Its immediate to verify the mean and speed of mean reversion reported in Proposition 6. We omit the steps to obtain the ergodic variance of $\tilde{B}_t$ and the moments for $\tilde{Z}_t$ as they are same as in the proof of Proposition 3.

A.5 Proof of Proposition 7

We prove the case when s is i.i.d. The details of the more general case when s is Markov are in an online appendix. The first order conditions of problem (42) with $\mathcal{B}$ and τ yield

$$\begin{aligned}
V'(B_)&=\frac{\mathbb{E}[V'(\mathcal{B}(s))U_c(s)p(s)]}{\mathbb{E}[U_c(s)p(s)]},\\
\frac{\partial U}{\partial\tau}(\tau(s),s)-\mu(s)\frac{\partial\mathcal{X}}{\partial\tau}(\tau(s),s)&-\frac{\partial U_c}{\partial\tau}(\tau(s),s)\frac{\mathcal{B}_}{\beta\mathbb{E}[U_cp]}(\mu(s)-\mu_)=0. \qquad (93)
\end{aligned}$$

where $\mu\_=\beta V'(\mathcal{B}\_)$. Let $\boldsymbol{\epsilon}=(\epsilon_g,\epsilon_p,\epsilon_\theta)$ be the mean zero innovations to the stochastic processes (g,p,θ) and $\xi\equiv\mathbb{E}[U_cp]$. Equation (93) implicitly defines the function $\tau(\mu,\mu\_,\xi,\mathcal{B},\sigma\boldsymbol{\epsilon})$ and using $\tau(\cdot)$ we can define functions $\mathcal{U}_c(\mu,\mu\_,\xi,\mathcal{B},\sigma\boldsymbol{\epsilon})\equiv U_c(\tilde{\mu},\tilde{\mu}\_,\tilde{\xi},\mathcal{B}\_,\sigma\boldsymbol{\epsilon})(1+\sigma\epsilon_p)$ and $\mathcal{X}(\mu,\mu\_,\xi,\mathcal{B},\sigma\boldsymbol{\epsilon})$. The optimal policy functions $\tilde{\mu}(\mathcal{B}\_;\sigma)$, $\tilde{\xi}(\mathcal{B}\_;\sigma)$ and $\tilde{\mathcal{B}}(\sigma\boldsymbol{\epsilon},\mathcal{B}\_;\sigma)$ satisfy

$$\frac{\mathcal{B}\_\mathcal{U}_c(\tilde{\mu},\tilde{\mu}\_,\tilde{\xi},\mathcal{B}\_,\sigma\boldsymbol{\epsilon})}{\beta\tilde{\xi}}+\mathcal{X}(\tilde{\mu},\tilde{\mu}\_,\tilde{\xi},\mathcal{B}\_,\sigma\boldsymbol{\epsilon})=\tilde{\mathcal{B}}, \qquad (94)$$
$$0=\mathbb{E}\left[(\tilde{\mu}-\tilde{\mu}\_)\mathcal{U}_c\right], \qquad (95)$$
$$\xi=\tilde{\mathbb{E}}[\mathcal{U}_c]. \qquad (96)$$

For brevity we have dropped the arguments of the $\tilde{}$ functions and $\tilde{\mu}$ represents $\tilde{\mu}(\tilde{B}(\sigma\boldsymbol{\epsilon},\mathcal{B}\_;\sigma);\sigma)$ while $\tilde{\mu}\_$ represents $\tilde{\mu}(\mathcal{B}\_;\sigma)$. The steady state associated with $\mathcal{B}\_$ and after setting $\sigma=0$ has $\mu(s)=\mu\_=\bar{\mu}$.[32]

As in proof of Proposition 3, we now proceed to compute the terms needed for the second order Taylor expansion of $\tilde{\mathcal{B}}(\sigma\boldsymbol{\epsilon},\mathcal{B}\_;\sigma)$ at $\sigma=0$. At the outset we derive some properties of $\tau(\cdot)$ that will be useful in simplifying terms as we go along. Differentiating (93) with respect to $\mathcal{B},\xi$ and μ and evaluating them at $\sigma=0$ where $\mu(s)=\mu\_$, we obtain that the following derivatives are zero:

$$\frac{\partial\tau}{\partial\mathcal{B}},\frac{\partial\tau}{\partial\xi},\frac{\partial^2\tau}{\partial\mathcal{B}^2},\frac{\partial^2\tau}{\partial\xi^2},\text{ and }\frac{\partial^2\tau}{\partial\mathcal{B}\partial\xi}.$$

[32]In particular for a given $\mathcal{B}\_$ the steady state $\bar{\mu}$, $\bar{\xi}$ solve

$$\frac{\mathcal{B}\_(1-\beta)}{\beta}=\mathcal{X}(\bar{\mu},\bar{\mu},\bar{\xi},\mathcal{B}\_,0)$$
$$\bar{\xi}=\mathcal{U}_c(\bar{\mu},\bar{\mu},\bar{\xi},\mathcal{B}\_,0)=\bar{\mathcal{U}}_c.$$

In addition,

$$\frac{\partial^2\tau}{\partial\mathcal{B}\partial\mu}+\frac{\partial^2\tau}{\partial\mathcal{B}\partial\mu\_}=0 \text{ and } \frac{\partial^2\tau}{\partial\xi\partial\mu}+\frac{\partial^2\tau}{\partial\xi\partial\mu\_}=0.$$

These properties extend to $\mathcal{U}_c(\mu,\mu\_,\xi,\mathcal{B},\sigma\boldsymbol{\epsilon})$ and $\mathcal{X}(\mu,\mu\_,\xi,\mathcal{B},\sigma\boldsymbol{\epsilon})$.

We begin by obtaining the first order terms. Differentiate (94)-(96) of w.r.t $\mathcal{B}\_$ to obtain

$$\tilde{B}_{\mathcal{B}\_}=1,\quad \tilde{\mu}_{\mathcal{B}\_}=\frac{1-\beta}{\beta(-\bar{\mathcal{X}}_{\mu}-\bar{\mathcal{X}}_{\mu\_})},\quad \tilde{\xi}_{\mathcal{B}\_}=\bar{\mathcal{U}}_{c,\mu}\tilde{\mu}_{\mathcal{B}\_}+\bar{U}_{c,\mu\_}\tilde{\mu}_{\mathcal{B}\_}. \tag{97}$$

Next differentiate with respect to $\boldsymbol{\epsilon}$ to obtain

$$\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}=\left(\frac{\mathcal{B}\_\bar{\mathcal{U}}_{c,\boldsymbol{\epsilon}}}{\bar{\xi}\beta}+\bar{\mathcal{X}}_{\boldsymbol{\epsilon}}\right)\left(\frac{1}{\beta}+\bar{\mathcal{X}}_{\mu\_}\tilde{\mu}_{\mathcal{B}\_}-\frac{\mathcal{B}\_\bar{\mathcal{U}}_{c,\mu}}{\beta\bar{\xi}}\tilde{\mu}_{\mathcal{B}\_}\right)^{-1}.$$

Using $\tilde{\mu}_{\mathcal{B}\_}=\mathcal{O}(1-\beta)$ from (97) it follows that

$$\left(\frac{1}{\beta}+\bar{\mathcal{X}}_{\mu\_}\tilde{\mu}_{\mathcal{B}\_}-\frac{\mathcal{B}\_\bar{\mathcal{U}}_{c,\mu}}{\beta\bar{\xi}}\tilde{\mu}_{\mathcal{B}\_}\right)^{-1}=\beta+\mathcal{O}(1-\beta)$$

and thus

$$\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}=\left(\frac{\mathcal{B}\_\bar{\mathcal{U}}_{c,\boldsymbol{\epsilon}}}{\bar{\mathcal{U}}_c}+\beta\bar{\mathcal{X}}_{\boldsymbol{\epsilon}}\right)+\mathcal{O}(1-\beta) \tag{98}$$

Finally differentiating with respect to σ and some algebra yields $\tilde{\mu}_{\sigma}=\tilde{\mathcal{B}}_{\sigma}=\tilde{\xi}_{\sigma}=0$.

Next we compute the the second order terms. Differentiating twice with respect to $\mathcal{B}\_$ and simplifying we find that

$$\tilde{\mathcal{B}}_{\mathcal{B}\_\mathcal{B}\_}=0,\quad \tilde{\mu}_{\mathcal{B}\_\mathcal{B}\_}=\frac{(\bar{\mathcal{X}}_{\mu\mu}+2\bar{\mathcal{X}}_{\mu\mu\_}+\bar{\mathcal{X}}_{\mu\_\mu\_})}{\bar{\mathcal{X}}_{\mu}+\bar{\mathcal{X}}_{\mu\_}}\tilde{\mu}^2_{\mathcal{B}\_}. \tag{99}$$

It is easy to show that the cross-partials with σ are zero. Next differentiatiating equation (95) twice with respect to σ we can show that

$$\tilde{B}_{\sigma\sigma}=-2\mathbb{E}\left[\left(\frac{\bar{\mathcal{U}}_{c,\boldsymbol{\epsilon}}}{\bar{\mathcal{U}}_c}\boldsymbol{\epsilon}\right)(\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}\boldsymbol{\epsilon})\right]-\mathbb{E}[\boldsymbol{\epsilon}'\tilde{\mathcal{B}}_{\boldsymbol{\epsilon\epsilon}}\boldsymbol{\epsilon}]+\mathcal{O}(1-\beta). \tag{100}$$

We are left with $\tilde{\mathcal{B}}_{\boldsymbol{\epsilon\epsilon}}$ which, in principle, can be quite complicated with risk aversion. However, we show that for computing the ergodic mean of effective debt and the speed of mean reversion we do not need the second derivatives with respect to $\boldsymbol{\epsilon}$. To see this note that the Taylor expansion of $\tilde{\mathcal{B}}$ after dropping terms that we have already shown to be zero is

$$\begin{aligned}\tilde{B}(\sigma\boldsymbol{\epsilon},\mathcal{B}\_;\sigma)=&\mathcal{B}\_+\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}\boldsymbol{\epsilon}+\frac{1}{2}\boldsymbol{\epsilon}'\tilde{\mathcal{B}}_{\boldsymbol{\epsilon\epsilon}}\boldsymbol{\epsilon}\sigma^2-\frac{1}{2}\mathbb{E}[\boldsymbol{\epsilon}'\tilde{\mathcal{B}}_{\boldsymbol{\epsilon\epsilon}}\boldsymbol{\epsilon}]\sigma^2\\&-\mathbb{E}\left[\left(\frac{\bar{\mathcal{U}}_{c,\boldsymbol{\epsilon}}}{\bar{\mathcal{U}}_c}\boldsymbol{\epsilon}\right)(\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}\boldsymbol{\epsilon})\right]\sigma^2+\mathcal{O}((1-\beta)\sigma^2,\sigma^3).\end{aligned}$$

Taking expectations we then have that

$$\mathbb{E}\left(\tilde{B}(\sigma\boldsymbol{\epsilon},\mathcal{B}\_;\sigma)|\mathcal{B}\_\right)=\mathcal{B}\_-\mathbb{E}\left[\left(\frac{\bar{\mathcal{U}}_{c,\boldsymbol{\epsilon}}}{\bar{\mathcal{U}}_c}\boldsymbol{\epsilon}\right)(\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}\boldsymbol{\epsilon})\right]\sigma^2+\mathcal{O}((1-\beta)\sigma^2,\sigma^3). \tag{101}$$

Now substituting for $\tilde{\mathcal{B}}_{\boldsymbol{\epsilon}}$ from equation (98) in equation (101) and noting that for any $\mathcal{B}$ we have

$$\mathcal{R}_{\tau(\mathcal{B})}(\sigma\boldsymbol{\epsilon};\sigma) - \mathcal{R}_{\tau(\mathcal{B})}(0;0) = \frac{\partial \mathcal{R}_{\tau(\mathcal{B})}(\sigma\boldsymbol{\epsilon};\sigma)}{\partial \boldsymbol{\epsilon}}\sigma\epsilon + \mathcal{O}(\sigma^2),$$
$$\chi_{\tau(\mathcal{B})}(\sigma\boldsymbol{\epsilon};\sigma) - \chi_{\tau(\mathcal{B})}(0;0) = \frac{\partial \chi_{\tau(\mathcal{B})}(\sigma\boldsymbol{\epsilon};\sigma)}{\partial \boldsymbol{\epsilon}}\sigma\epsilon + \mathcal{O}(\sigma^2),$$

we get

$$\mathbb{E}\left(\tilde{B}(\sigma\boldsymbol{\epsilon},\mathcal{B}_{\_};\sigma)|\mathcal{B}_{\_}\right) = \mathcal{B}_{\_} - \beta^2\text{var}\left(\mathcal{R}_{\tau(\mathcal{B}_{\_})}\right)\mathcal{B}_{\_} - \beta^2\text{cov}\left(\mathcal{R}_{\tau(\mathcal{B}_{\_})},\chi_{\tau(\mathcal{B}_{\_})}\right) + \mathcal{O}((1-\beta)\sigma^2,\sigma^3). \tag{102}$$

As $\tau(\mathcal{B})$ satisfies the implicit equation (45), we can differentiate (45) with respect to $\mathcal{B}$ to see that

$$\frac{\partial\tau(\mathcal{B})}{\partial\mathcal{B}}\mathbb{E}\frac{\partial}{\partial\tau}\mathcal{X}_{\tau_{\mathcal{B}}} = \frac{1-\beta}{\beta} = \mathcal{O}(1-\beta).$$

As $\beta\to1$, $\tau(\mathcal{B})\to\tau(0)$ and thus $\mathbb{E}\left(\frac{\partial\mathcal{X}_{\tau_{\mathcal{B}}}}{\partial\tau}\right)$ limits to a constant as $\beta\to1$ which implies that

$$\frac{\partial\tau(\mathcal{B})}{\partial\mathcal{B}},\ \frac{\partial}{\partial\mathcal{B}}\mathcal{X}_{\tau(\mathcal{B})} \text{ and } \frac{\partial}{\partial\mathcal{B}}\mathcal{R}_{\tau(\mathcal{B})} = \mathcal{O}(1-\beta).$$

After taking fist order conditions with respect to $\mathcal{B}$ of the variance minimization problem (44) we find that $\mathcal{B}^*$ satisfies

$$\mathbb{E}\left(\mathcal{R}_{\tau(\mathcal{B}^*)}\mathcal{B}^* + \mathcal{X}_{\tau(\mathcal{B}^*)}\right)\left(\mathcal{R}_{\tau(\mathcal{B}^*)} + \mathcal{B}^*\frac{\partial}{\partial\mathcal{B}}\mathcal{R}_{\tau(\mathcal{B}^*)} - \frac{\partial}{\partial\mathcal{B}}\mathcal{X}_{\tau(\mathcal{B}^*)}\right) = 0.$$

Rearranging terms and using our expressions from above[33] we get

$$\mathcal{B}^* = -\frac{\text{cov}\left(\mathcal{R}_{\tau(\mathcal{B}^*)},\mathcal{X}_{\tau(\mathcal{B}^*)}\right)}{\text{var}\left(\mathcal{R}_{\tau(\mathcal{B}^*)}\right)} + \mathcal{O}\left(1-\beta\right).$$

Finally $\tau(\mathcal{B}^*) - \tau(\mathcal{B}) = \mathcal{O}(1-\beta)$ and thus

$$\mathcal{B}^* = -\frac{\text{cov}\left(\mathcal{R}_{\tau(\mathcal{B}^*)},\mathcal{X}_{\tau(\mathcal{B}^*)}\right)}{\text{var}\left(\mathcal{R}_{\tau(\mathcal{B}^*)}\right)} = -\frac{\text{cov}\left(\mathcal{R}_{\tau(\mathcal{B})},\mathcal{X}_{\tau(\mathcal{B})}\right)}{\text{var}\left(\mathcal{R}_{\tau(\mathcal{B})}\right)} + \mathcal{O}(1-\beta). \tag{103}$$

Using (103) and $1 - x\sigma^2 = \frac{1}{1+x\sigma^2} + \mathcal{O}(\sigma^3)$ we then can rewrite (102) as

$$\mathbb{E}\left(\tilde{B}(\sigma\boldsymbol{\epsilon},\mathcal{B}_{\_};\sigma)|\mathcal{B}_{\_}\right) = \mathcal{B}^* + \frac{1}{1+\beta^2\text{var}\left(\mathcal{R}_{\tau(\mathcal{B}^*)}\right)}(\mathcal{B}_{\_} - \mathcal{B}^*) + \mathcal{O}((1-\beta)\sigma^2,\sigma^3).$$

The expressions for the ergodic mean of $\tilde{B}$ and speed of mean reversion reported in Proposition 7 follow immediately.

---

[33] $\mathcal{B}^*$ that solves the the risk minimizing level of debt that solves (44) is implicitly as a function of β. It can be verified that $\lim_{\beta\to1}\tau_{\mathcal{B}^*(\beta);\beta} = \tau(0)$. For the remainder of this proof we will drop the implicit dependence on β.

B Discount Factor Shocks

Quasilinear preferences

Following the specification in Albuquerque et al. (2016 forthcoming), we modify agent's preferences over consumption and labor in equation to

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \eta_t \left(c_t - \frac{l_t^{1+\gamma}}{1+\gamma} \right). \tag{104}$$

We maintain the rest of the assumptions of Section 3.3.

Proposition 8. *Suppose preferences satisfy (104). Let $\{\tilde{\tau}_t, \tilde{B}_t\}_t$ be the optimal tax, debt policy in an economy with initial debt B_{-1} and shocks (p, g, θ, η) with $\eta(s) = 1$ for all s. There exists an economy with shocks (p', g', θ', η'), where $p'(s) = 1$ and $\eta'(s) = p(s), g'(s) = \frac{g(s)}{p(s)}, \theta'(s) = \theta(s)p(s)^{-\left(\frac{\gamma}{1+\gamma}\right)}$ and initial debt $B'_{-1} = \eta'_{-1}B_{-1}$ such that the optimal policy $\{\tilde{\tau}'_t, \tilde{B}'_t\}_t$ for such economy satisfies $\tilde{\tau}'_t(s^t) = \tilde{\tau}_t(s^t)$ and $\tilde{B}'_t(s^t) = \eta'_t(\tilde{s^t})B(s^t)$ for histories s^t.*

Proof. Allowing for η_t modifies the $R_t = \frac{p_t \eta_{t-1}}{\beta \mathbb{E}_{t-1}\eta_t p_t}$. Define $\mathcal{B}_t = \eta_t B_t$ and multiplying the implementability constraint (9) by η_t, the optimal allocation maximizes (104) subject to

$$\frac{p_t \eta_t}{\beta \mathbb{E}_{t-1}\eta_t p_t}\mathcal{B}_{t-1} + \eta_t g_t = \eta_t(\theta_t l_t - l_t^{1+\gamma}) + \mathcal{B}_t, \tag{105a}$$

$$c_t + g_t = \theta_t l_t. \tag{105b}$$

Let μ_t, ξ_t be the multipliers on (105a) and (105b). The first order conditions with respect to c_t and l_t are

$$\xi_t = \eta_t,$$

$$-\eta_t l_t^{\gamma} + \mu_t \eta_t (\theta_t - (1+\gamma) l_t^{\gamma}) + \theta_t \xi_t = 0.$$

Substituting for ξ_t we obtain,

$$\tau_t l_t = \theta_t l_t - l_t^{1+\gamma} = \theta_t^{\frac{1+\gamma}{\gamma}} \Phi(\mu_t), \tag{106}$$

where $\Phi(\cdot)$ is defined in equation (58). Combining (106) with (105a) the optimal allocation satisfies

$$\frac{p_t \eta_t}{\beta \mathbb{E}_{t-1}\eta_t p_t}\mathcal{B}_{t-1} + \eta_t g_t = \eta_t \theta_t^{\frac{1+\gamma}{\gamma}} \Phi(\mu_t) + \mathcal{B}_t, \tag{107}$$

$$\mu_{t-1} = \mathbb{E}_{t-1}\eta_t p_t \mu_t. \tag{108}$$

The solutions $\{\mathcal{B}_t, \mu_t\}$ that solve (107) and (108) are invariant across economies that differ in $\{p_t, g_t, \theta_t, \eta_t\}_t$ as long they have the same values for $p_t\eta_t$, $\eta_t g_t$ and $\eta_t \theta_t^{\frac{1+\gamma}{\gamma}}$. □

Risk aversion

We show that the level of risk minimizing debt upto $\mathcal{O}(\sigma, (1-\beta))$ terms is the same if we use discount factor shocks $\eta(s)$ in place of payoff shocks $p(s)$. The counterpart of problem (42) with discount factor shocks can be also expressed recursively with the state variable $\mathcal{B}_t = \eta_t U_{c,t} B_t$. The implementability constraint then becomes

$$\frac{\mathcal{B}\_U_c(s)\eta(s)}{\beta \mathbb{E} U_c(s)\eta(s)} + \eta(s)\mathcal{X}(s) = \mathcal{B}(s).$$

In comparison to the same expression with payoff shocks

$$\frac{\mathcal{B}\_U_c(s)p(s)}{\beta \mathbb{E} U_c(s)p(s)} + \mathcal{X}(s) = \mathcal{B}(s),$$

both $\eta(s)$ and $p(s)$ appear in a similar fashion in the effective returns term. However, the deficits are adjusted by the discount factor shocks and not by the payoff shocks. Let $\mathcal{R}_{\mathrm{t}} = \frac{\eta_{\mathrm{t}} \mathrm{U}_{\mathrm{c,t}}}{\beta \mathbb{E}_{\mathrm{t}-1} \eta_{\mathrm{t}} \mathrm{U}_{\mathrm{c,t}}}$ and $\mathcal{X}_t = U_{c,t}c_t + U_{l,t}l_t$. Augmenting problem (44), the risk minimizing level $\hat{\mathcal{B}}^*$ with discount factor shocks is then given by

$$\hat{\mathcal{B}}^* \equiv \arg\min_{\mathcal{B}} \operatorname{var}\left(\mathcal{R}_{\tau(\mathcal{B})} B + PV(\eta \mathcal{X}_{\tau(\mathcal{B})})\right). \tag{109}$$

Proposition 9. *Let $\mathcal{B}^*$ and $\hat{\mathcal{B}}^*$ be solutions to minimization problems (44) and (109) respectively, we have*

$$\hat{\mathcal{B}}^* = \mathcal{B}^* + \mathcal{O}(\sigma, 1-\beta).$$

Proof. Arguments on the lines in Appendix A.5 shows that

$$\hat{\mathcal{B}}^* = \arg\min_{\mathcal{B}} \operatorname{var}\left[\left(\mathcal{R}_{\tau(0)}\mathcal{B} + PV(\eta \mathcal{X}_{\tau(0)})\right)\right] + \mathcal{O}(1-\beta).$$

The first order condition with respect to $\mathcal{B}$ gives

$$\mathbb{E}\left[\left(\mathcal{R}_{\tau(0)} - \frac{1}{\beta}\right)\left(\left(R_{\tau(0)} - \frac{1}{\beta}\right)\hat{\mathcal{B}}^* + PV(\eta \mathcal{X}_{\tau(0)})\right)\right] = 0.$$

A little algebra shows that, under the normalization of $\mathbb{E}\eta = 1$,

$$PV(\eta \mathcal{X}_{\tau(0)}) = PV(\mathcal{X}_{\tau(0)}) + PV(\eta)\mathbb{E}\mathcal{X}_{\tau(0)} + \mathcal{O}(\sigma^2) = PV(\mathcal{X}_{\tau(0)}) + \mathcal{O}(\sigma^2, (1-\beta)\sigma).$$

Thus

$$\operatorname{var}(R_{\tau(0)})\hat{\mathcal{B}}^* + \operatorname{cov}(\mathcal{R}_{\tau(0)}, PV(\mathcal{X}_{\tau(0)})) + \mathcal{O}(\sigma^3, (1-\beta)\sigma^2) = 0,$$

and after dividing by $\operatorname{var}(R_{\tau(0)})$ we recover

$$\hat{\mathcal{B}}^* = -\frac{\operatorname{cov}(\mathcal{R}_{\tau(0)}, PV(\mathcal{X}_{\tau(0)}))}{\operatorname{var}(R_{\tau(0)})} + \mathcal{O}(\sigma, 1-\beta).$$

□

C Numerical Appendix

To obtain the optimal allocation we approximate functions

$$\left\{c_{s,s\_}(\cdot), l_{s,s\_}(\cdot), \xi_{s,s\_}(\cdot), \mu'_{s,s\_}, \mathcal{B}_{s\_}(\cdot)\right\}$$

for $s, s\_ \in S \times S$, all defined on a compact interval $[\mu_{min}, \mu_{max}]$, that satisfy the following set of equations

$$0 = \begin{cases} \theta(s) l_{s,s\_}(\mu) - c_{s,s\_}(\mu) - g(s) \\ \frac{\mathcal{B}_{s\_}(\mu) p(s) U_c(c_{s,s\_}(\mu))}{\beta \mathbb{E}_{s\_} U_c(c(\mu)) p} - U_c(c_{s,s\_}(\mu)) c_{s,s\_}(\mu) - U_l(l_{s,s\_}(\mu)) l_{s,s\_}(\mu) - \mathcal{B}_s\left(\tilde{\mu}_{s,s\_}\right) \\ \tilde{\mu}_{s,s\_} - \min\left\{\mu_{max}, \max\left\{\mu'_{s,s\_}(\mu), \mu_{min}\right\}\right\} \\ U_l(l_{s,s\_}(\mu)) - \mu'_{s,s\_}(\mu)\left(U_l(l_{s,s\_}(\mu)) l_{s,s\_}(\mu) + U_l(l_{s,s\_}(\mu))\right) + \theta \xi_{s,s\_}(\mu) \\ \mu - \mathbb{E}_{s\_}\left[\mu'_{s,s\_}(\mu)\left(\frac{p(s) U_c(c_{s,s\_}(\mu))}{\mathbb{E}_{s\_} U_c(c(\mu)) p}\right)\right] \\ U_c(c_{s,s\_}(\mu)) - \mu'_{s,s\_}(\mu)\left(U_{cc}(c_{s,s\_}(\mu)) c_{s,s\_}(\mu) + U_c(c_{s,s\_}(\mu))\right) + \frac{\mathcal{B}_{s\_}(\mu) p(s) U_{cc}(c_{s,s\_}(\mu))}{\beta \mathbb{E}_{s\_} U_c(c(\mu)) p}\left(\mu'_{s,s\_}(\mu) - \mu\right) - \xi_{s,s\_}(\mu) \end{cases}$$

These equations are the resource constraints, implementability constraints, and the first order necessary conditions with respect to $l(s)$, $\mathcal{B}(s)$ and $c(s)$ for the $t \geq 1$ recursive Bellman equation stated in (42). We describe our numerical procedure below:

1. Domain: The functions are defined on domain that has a mix of discrete states, i.e., $s, s\_$ and continuous states $\mu \in [\mu_{min}, \mu_{max}]$.

   (a) To obtain S, and $\pi(s|s\_)$, we follow Kopecky and Suen (2010) to obtain a 5 state discrete approximation for the shock processes. For the shocks ϵ_p, ϵ_g we use a 5 state Gaussian quadrature.

   (b) The bounds on domain for μ indirectly impose bounds on $\mathcal{B}$. [34] We use 30 points on the μ grid, with more points near both the end points.

2. All functions are indexed with $s, s\_$ and for the μ dimension we use cubic splines as basis functions with the knot points placed on the 30 grid points that we chose in the previous step.

3. Next we iterate on the functions $\{\mathcal{B}_s(\cdot)\}_{s \in S}$

   (a) Start with a guess for $\left\{\mathcal{B}_s^j(\cdot)\right\}_{s \in S}$

   (b) For each point $\mu, s\_$ we use the $\mathcal{B}_s^j$ for evaluating $\mathcal{B}_s\left(\tilde{\mu}_{s,s\_}\right)$ with $\tilde{\mu}_{s,s\_} = \min\left\{\mu_{max}, \max\left\{\mu'_{s,s\_}(\mu), \mu_{min}\right\}\right\}$ in first equation in the system of equations above. Then using a non-linear root solver get $c(s;\mu,s\_), l(s;\mu,s\_), \xi(s;\mu,s\_), \mu'(s;\mu,s\_), \mathcal{B}(\mu,s\_)$ as $4S+1$ unknowns in $4S+1$ equations.

   (c) Lastly we update $\left\{\mathcal{B}_s^{j+1}(\cdot)\right\}_s$ by interpolating the $\mathcal{B}(\mu, s)$ that we solved in the previous step.

[34] Since the natural debt limit implies that a $\mu_{max} = \infty$, our μ grid, $[\mu_{min}, \mu_{max}]$ ensures that the implied debt limits are tighter than the natural debt limit

(d) Iterate until $\sup_{\mu,s\_} \|\mathcal{B}_{s\_}^{j+1}(\mu) - \mathcal{B}_{s\_}^{j}(\mu)\| \leq 1e-5$

4. To check that our solution to the first order conditions is a global maximum, we compute the value function for the government using the policy rules solved in steps 1-4 and use a numerical global optimizer to confirm that the solution to the first order condition also achieve the maximum value for the government.

5. The time-0 problem is solved using the time 0 first order conditions and our approximation of the solution $t \geq 1$ policy rules.

Competitive Equilibrium

Here we detail the methodology used to solve for the competitive equilibrium under the tax rule in equation (51). Given the tax rule in equation (51)[35] a competitive equilibrium satisfies

$$c_t^{-\alpha}(1-\tau_t)\theta(s_t) = l_t^{\gamma},$$

$$c_{t-1}^{-\alpha} = \beta \mathbb{E}_{t-1} \frac{p(s_t)}{q_{t-1}} c_t^{\alpha},$$

$$c_t + g(s_t) = \theta(s_t) l_t.$$

To compute the competitive equilibrium we begin by noting that, conditional on the shocks $g(s_t), \theta(s_t)$, the combination of labor leisure choice and the aggregate resource constraint pin down c and l for any choice of τ. This implicitly defines the functions $c(\tau, s)$ and $l(\tau, s)$. The lagged variables in equation (51) imply that the state variable for this problem will be at least $z_{t-1} = (\tau_{t-1}, R_{t-1}, B_{t-1}, s_{t-1})$.[36] Our goal is to determine the $\tau(s_t)$ that satisfy (51) conditional on z_{t-1}, however we face the immediate problem that many of the variables on the R.H.S. of (51) are endogenous objects. Fortunately this reduces to solving a system of S non-linear equations and S unknowns.

Begin with a guess of the equilibrium $\tau(s)$ given the previous state $z\_$. Output for each realization of the aggregate shock s is given by $y(s) = \theta(s) l(\tau(s), s)$, while the return is

$$R(s) = \frac{(y\_ - g(s\_))^{-\alpha} p(s)}{\beta \sum_s \Pi(s, s\_) c(\tau(s), s)^{-\alpha}}.$$

From equation (51) the prescribed tax policy can be computed as

$$\hat{\tau}(s) = (1-\rho_\tau)\overline{\tau} + \rho_\tau \tau\_ + \rho_Y \log y(s) + \rho_{Y\_} \log y\_ + \rho_g g(s) + \rho_{g\_} g(s\_) + \rho_R R(s) + \rho_{R\_} R\_ + \rho_{B\_} \log B\_.$$

The competitive equilibrium is found by solving for the vector $\tau(s)$ such that $\hat{\tau}(s) = \tau(s)$ for all s. For a given realization of the shock s the current state z is determined by using the government budget constraint to solve for $B(s)$ (as $y(s)$ and $R(s)$ were determined above).

[35] We additionally impose that government debt to GDP cannot exit the region [.1, 3.]. The government will only violate (51) if its debt level would exit this region and would set tax policy to ensure issued government debt was set at the boundary of this admissible region. This event never occurs in our simulations and the numerical procedure to do this follows identically to what is described below.

[36] y_{t-1} is not required as τ_{t-1} and s_{t-1} jointly determine y_{t-1}.

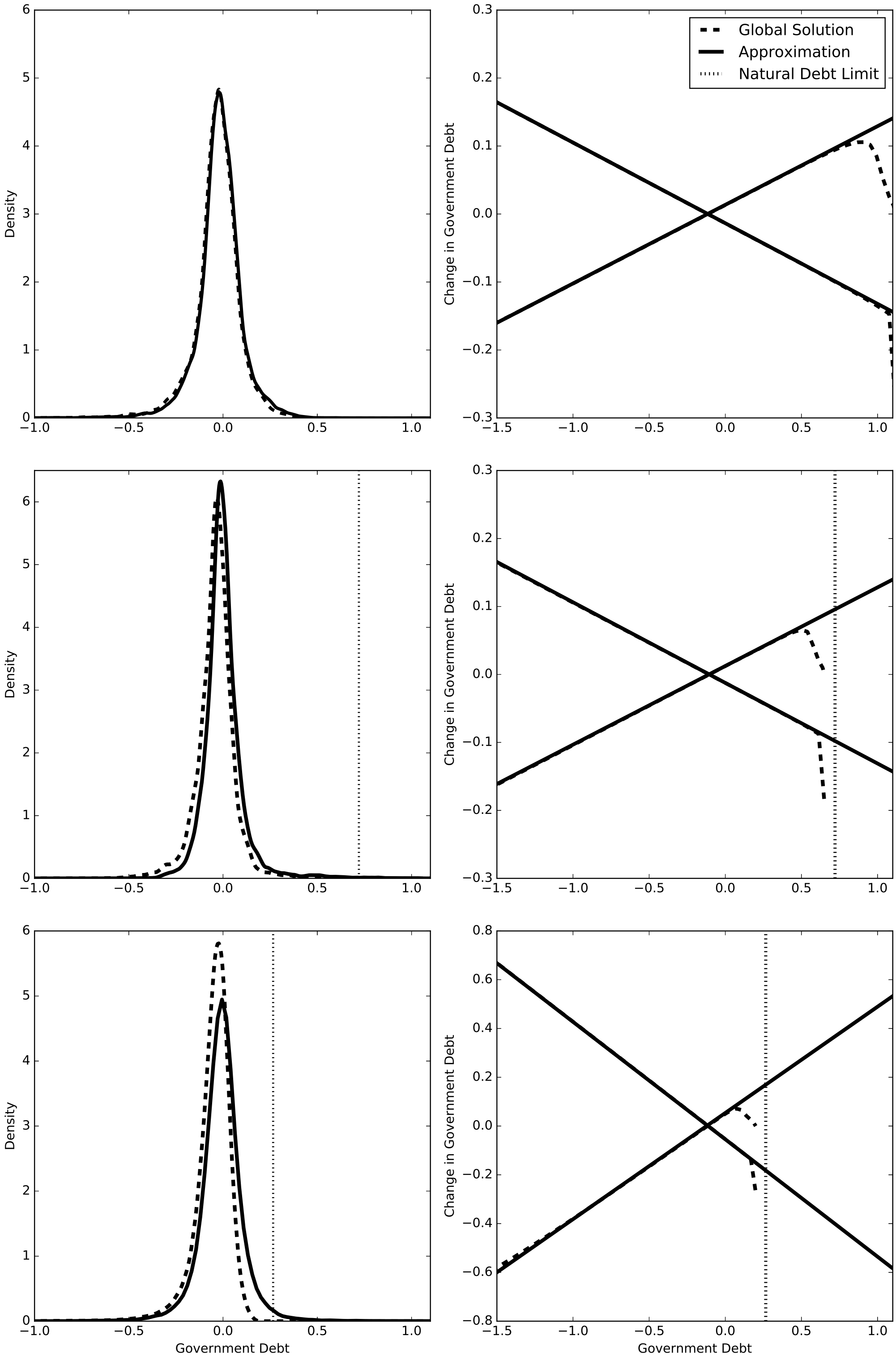

Figure I: Using the quadratic approximation (red line) and a more accurate global approximation (black line), the top, middle, and bottom panels plot smoothed kernel densities (left side) and decision rules (right side) associated with baseline parameters in Table I, high discount factor ($\beta = 0.90$) and large shocks ($\sigma = 4$) settings. The right panel displays policies $\tilde{B}(s, B\_) - B\_$ for two values of s that correspond to the smallest and the largest pairs of $(g(s), p(s))$.

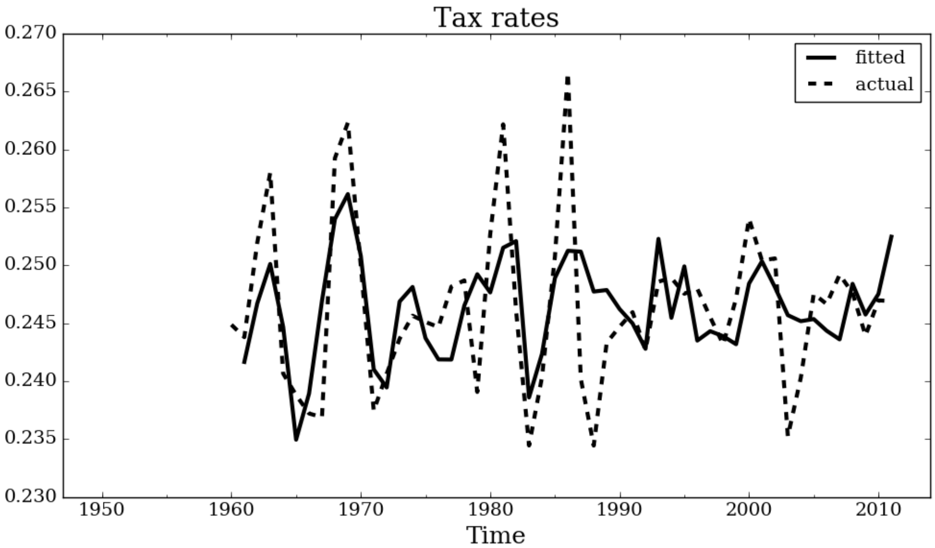

Figure II: Fitted debt versus (H.P. filtered) average marginal tax rates

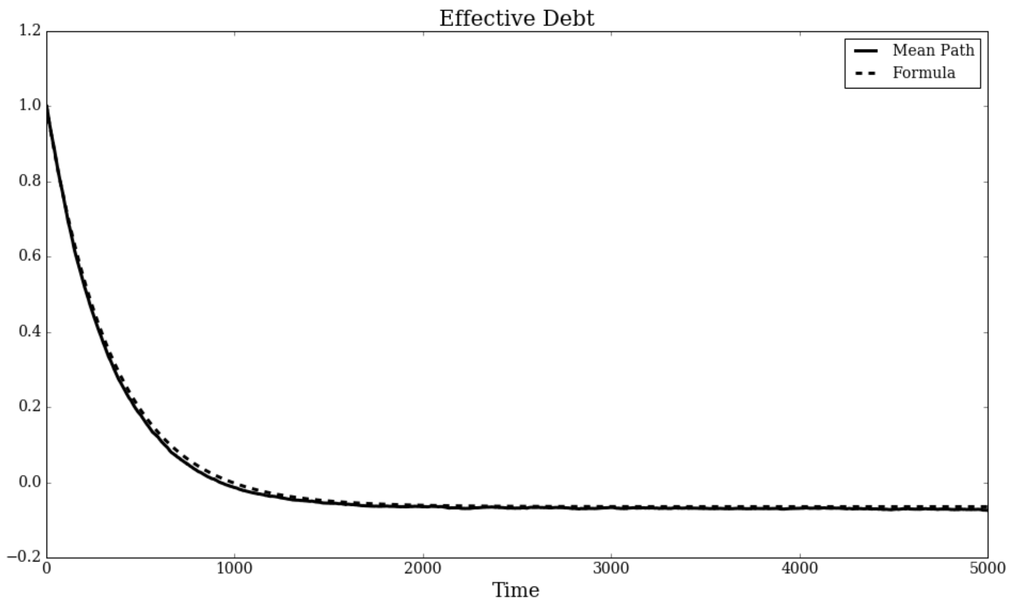

Figure III: The solid line is the conditional mean path for effective debt, $\mathbb{E}_0\mathcal{B}_t$ after averaging across 10000 simulated paths. The dashed line is computed using the formula 52.

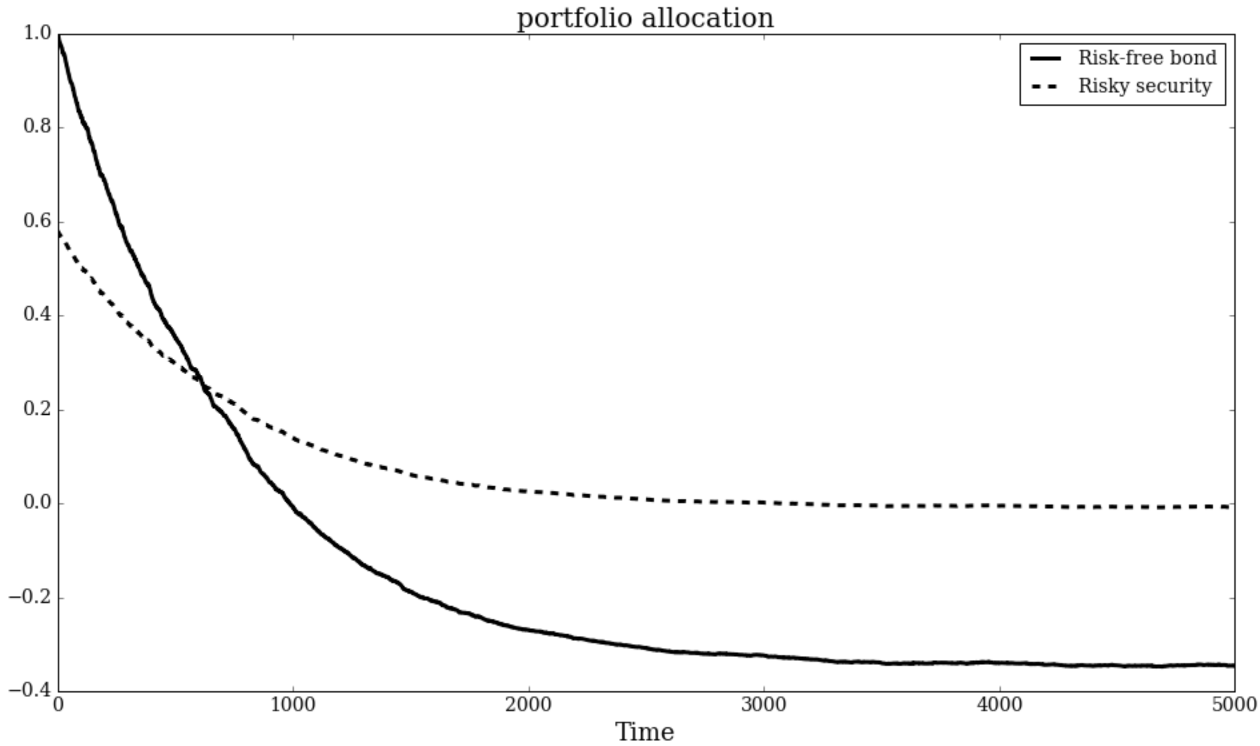

Figure IV: Marginal utility weighted holdings of the risk free bond (solid line) and "stock market" security (dashed line). Negative values implies that the government is shorting the security.

Parameter	Value	Moments	Values
$\bar{g}$	0.25	mean government expenditure relative to output	25%
std. ϵ_g	0.01	std. of log government expenditures	2.6%
std. $\epsilon_{\hat{p}}$	0.05	std. of returns of debt portfolio	5.1%
χ	0.6	correlation of returns and log government expenditures	0.078

Table I: Parameters and moments used for comparing the accuracy of the quadratic approximations in the quasilinear economy.

Param	Value	Moment	Model	Data
		Log Output		
σ_θ	0.02	std. dev	1.7%	1.6%
ρ_θ	0.35	auto corr	0.23	0.23
		Returns		
σ_p	0.05	std. dev	5.2%	5.1%
χ_p	0.25	corr with $\log y_t$	-0.008	-0.004
		Log expenditures		
$\bar{g}$	0.25	mean g_t/y_t	25%	25%
σ_g	0.02	std. dev	2.6%	2.6%
χ_g	-0.2	corr with $\log y_t$	-0.14	-0.15

Table II: Parameters and targeted moments in the competitive equilibrium with fitted U.S. tax policies.

Parameter	Value
$\bar{\tau}$	0.25 (0.021)
$\rho_{\tau\_}$	0.19 (0.14)
$\rho_y, \rho_{y\_}$	0.09 (0.08),0.21 (0.08)
$\rho_g, \rho_{g\_}$	0.11 (0.06), 0.11 (0.06)
$\rho_R, \rho_{R\_}$	0.04 (0.03), -0.02 (0.03)
$\rho_{B\_}$	0.02, (0.05)

Table III: OLS estimates for tax rule. The numbers in brackets are standard errors.

Moments		Baseline		1 yr. yields		Risk-free bond	
	global solution	quadratic approx.	VAR	global solution	quadratic approx.	global solution	quadratic approx.
Effective debt, $\mathcal{B}_t$							
mean	-7%	-7%	-6%	-24%	-23%	-42%	-42%
half-life (years)	237	244	249	655	678	1244	1299
std.	18%	20%	-	25%	33%	18%	46%
Tax rates, Z_t							
mean	20%	20%	-	20%	20%	20%	20%
half-life (years)	263	244	-	667	678	1234	1299
std.	0.2%	0.4%	-	0.3%	0.7%	0.2%	0.9%

Table IV: Ergodic moments for effective debt and tax revenues.

Parameter	Value
α_y	-0.83
α_χ	0.50
$\mathrm{cov}(\mathcal{R}, \log y)/\mathrm{var}(\mathcal{R})$	0.063
$\mathrm{cov}(\mathcal{R}, \chi)/\mathrm{var}(\mathcal{R})$	-0.006
$\mathrm{var}(\mathcal{R})$	0.003

Table V: VAR estimates

Portfolio holdings	global solution	quadratic approx
risk-free bond	-42.40%	-43.00%
risky asset	-0.05%	0.06%

Table VI: Ergodic portfolio using global solution and formula (48)