

## Optimal Fiscal Policy with Recursive Preferences

Anastasios G. Karantounias

Working Paper 2013-7c

September 2013 (Revised January 2018)

**Abstract:** I study the implications of recursive utility, a popular preference specification in macrofinance, for the design of optimal fiscal policy. Standard Ramsey tax-smoothing prescriptions are substantially altered. The planner overinsures by taxing less in bad times and more in good times, mitigating the effects of shocks. At the intertemporal margin, there is a novel incentive for introducing distortions that can lead to an ex-ante capital subsidy. Overall, optimal policy calls for a much stronger use of debt returns as a fiscal absorber, leading to the conclusion that actual fiscal policy is even worse than we thought.

JEL classification: D80, E62, H21, H63

Key words: Ramsey plan, tax smoothing, Epstein-Zin, recursive utility, excess burden, labor tax, capital tax, martingale, fiscal insurance

---

The author is grateful to the editor (Philipp Kircher) and to three anonymous referees for insightful comments that improved the final product. He also thanks Roc Armenter, David Backus, Pierpaolo Benigno, Dominique Brabant, R. Anton Braun, Vasco Carvalho, Lawrence Christiano, Lukasz Drozd, Kristopher S. Gerardi, Mikhail Golosov, Jonathan Halket, Lars Peter Hansen, Karen Kopecky, Hanno Lustig, Juan Pablo Nicolini, Demian Pouzo, Victor Rios-Rull, Richard Rogerson, Thomas J. Sargent, Yongseok Shin, and Stanley E. Zin. He also thanks Christopher Sleet for his discussion, seminar participants at the Bank of Portugal, Carnegie Mellon University, the Einaudi Institute of Economics and Finance, the European University Institute, LUISS Guido Carli University, Northwestern University, the University of California at Davis, the University of Hong Kong, the University of Notre Dame, the University of Oxford, Universitat Autònoma de Barcelona, Universitat Pompeu Fabra, the University of Reading, conference participants at the first New York University Alumni Conference, the CRETE 2011 Conference, the 2012 Society for Economic Dynamics meetings, the 2012 European Economic Association annual congress, the 2013 American Economic Association meetings, and the Federal Reserve Banks of Atlanta, Chicago, Minneapolis, and St. Louis. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Anastasios Karantounias, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309, [anastasios.karantounias@atl.frb.org](mailto:anastasios.karantounias@atl.frb.org).

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at [www.frbatlanta.org](http://www.frbatlanta.org). Click "Publications" and then "Working Papers." To receive e-mail notifications about new papers, use [frbatlanta.org/forms/subscribe](http://frbatlanta.org/forms/subscribe).

# 1 Introduction

The basic fiscal policy prescription in dynamic, stochastic, frictionless economies is *tax-smoothing*. Labor taxes should be essentially constant and any kind of shock should be absorbed by proper debt management. This result comes from the seminal work of [Lucas and Stokey \(1983\)](#) and [Chari et al. \(1994\)](#) and forms the heart of dynamic Ramsey policy.

In this paper, I show that if we differentiate between risk aversion and intertemporal elasticity of substitution and use the recursive preferences of [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#), the conventional normative tax-smoothing result breaks down. Optimal policy generates large surpluses and deficits by prescribing high taxes in good times and low taxes in bad times. Furthermore, in contrast to standard Ramsey results, labor taxes are persistent independent of the stochastic properties of exogenous shocks and capital income should be subsidized.

The coefficients of intertemporal elasticity of substitution and risk aversion are two parameters that are a priori important in shaping dynamic policy. They control the desirability of taxing in the current versus future periods and the aversion towards shocks that hit the government budget. Unfortunately, time-additive expected utility renders the analysis of the implications of these two parameters on optimal policy impossible. Moreover, since the temporal dimension of risk is ignored, questions about the implications of long-run fiscal risks on current tax and debt policies can be answered only in a limited way.

More crucially, optimal fiscal policy revolves around the proper choice of taxes and government securities to maximize welfare. To determine the desirability of debt securities, a plausible model of returns is needed. Conventional Ramsey analysis uses time-additive expected utility, a specification which is notorious for its difficulty in generating realistic asset prices, casting therefore doubts on the merits of standard tax-smoothing prescriptions. The failure to match risk premia has made the empirically more successful recursive preferences the norm in the literature that merges macroeconomics and finance.<sup>1</sup> It is natural to speculate that any model that prices better risk will alter the qualitative and quantitative nature of fiscal policy. However, little is known about recursive utility and optimal fiscal policy even in the simplest Ramsey setup. This is the task of the current paper.

Consider first an economy without capital as in [Lucas and Stokey \(1983\)](#). Linear taxes and state-contingent debt are used in order to finance an exogenous stream of stochastic government expenditures. A benevolent planner chooses under commitment the policy that maximizes the utility of the representative household. There are two basic results with time-additive expected utility: First, the labor tax should be *constant* if period utility features constant elasticities. Even when elasticities are not constant, the volatility of the labor tax is quite small. Second, whenever the labor tax varies, it inherits the stochastic properties of the exogenous shocks. Thus, optimal

---

<sup>1</sup>The literature is vast. See indicatively [Tallarini \(2000\)](#), [Bansal and Yaron \(2004\)](#), [Piazzesi and Schneider \(2007\)](#), [Hansen et al. \(2008\)](#), [Gourio \(2012\)](#), [Rudebusch and Swanson \(2012\)](#), [Petrosky-Nadeau et al. \(2013\)](#) and [Ai and Bansal \(2016\)](#) among others.

labor taxes do not constitute a distinct source of persistence in the economy.

As I argued earlier, both of these classic results are overturned in the same economy with recursive preferences. There is a simple, yet powerful intuition for that. Assume that risk aversion is greater than the inverse of the intertemporal elasticity of substitution. In that case, the household sacrifices smoothing over time in order to have a smoother consumption profile over states, becoming effectively *averse* to volatility in future utilities. As a response, the planner *attenuates* utility volatility by taxing less in bad times, offsetting therefore the effects of an adverse fiscal shock, and taxing more in good times, mitigating the benefits of a favorable fiscal shock.

What is the mechanism behind this intuition? The entire action is coming from the pricing of state-contingent claims with recursive utility. The planner hedges fiscal risk by issuing state-contingent debt against low spending shocks, to be paid by surpluses, and buys assets against high spending shocks, that are used to finance government deficits. With recursive utility the planner “over-insures,” that is, he sells more debt against low spending shocks relative to the expected utility benchmark. Consequently, taxes are higher in good times when spending is low, in order to repay the high levels of maturing debt. The reason for over-insurance is simple: by issuing more debt against good times the planner depresses future utilities. This reduction in utility is *priced* with recursive preferences, raising the stochastic discount factor. Thus, the price of state-contingent claims that the planner sells *rises*, making state-contingent debt against good times *cheaper*. So more revenue can be raised from debt issuance and the planner can relax the budget constraint, which is welfare-improving. Similarly, by purchasing more assets and taxing less against high spending shocks, the planner raises utility and therefore decreases the stochastic discount factor, relaxing again the government budget constraint.

Hence, the planner is trading off some tax volatility for more beneficial prices of state-contingent debt. The additional curvature of the utility function with respect to the “long-run”, as captured by future utilities, *amplifies* fiscal insurance, depressing ultimately risk premia. Optimal policy prescribes high returns for bond-holders when government spending is low, paid for with high taxes. In contrast, optimal policy prescribes capital losses for bond-holders when spending is high, allowing large deficits with low taxes. In fact, at high levels of government debt, the over-insurance efforts of the government can lead to a positive conditional covariance of the stochastic discount factor with the returns on the government debt portfolio, implying a *negative* conditional risk premium of government debt. The economics behind this remarkable result make sense: “good” times with low spending shocks can become “bad” times with very high tax rates. Thus, the household is happy to accept a negative premium for a risky security that pays well when distortionary taxes are high.

With recursive preferences a tax rate at a future period affects the entire sequence of one-period stochastic discount factors up to that period, due to the forward-looking nature of future utilities. As a result, the planner does not choose future tax rates independently from the past, but designs persistent policies in order to properly affect the entire sequence of prices of state-contingent claims.

Furthermore, it is cheaper on average to issue debt and postpone taxation, leading optimally to *back-loading* of tax distortions.

Recursive utility introduces non-trivial complications to the numerical analysis of the Ramsey problem. Value functions appear in the constraints since they affect the pricing of the government debt portfolio, hindering the contraction property, introducing non-convexities and complicating the calculation of the state space. A separate contribution of the paper is to deal with these issues and provide a numerical solution of the optimal taxation problem. In a series of numerical exercises I demonstrate the volatility and persistence of the tax rate and analyze the implications for the debt-to-output ratio.

As a final exercise, I quantify the *optimal* use of debt returns and tax revenues for the absorption of fiscal shocks and contrast it to the empirical findings of [Berndt et al. \(2012\)](#). [Berndt et al. \(2012\)](#) measure how fiscal shocks are absorbed by reductions in debt returns (the debt valuation channel or else *fiscal insurance*) or by increases in tax revenues (the surplus channel) in post-war U.S. data and find evidence of limited but non-negligible fiscal insurance. In contrast, optimal policy in an expected utility economy prescribes that the majority of fiscal risk should be absorbed by reductions in returns. Turning to a recursive utility economy, the debt valuation channel is even more prominent and can surpass 100%; fiscal insurance compensates for the fact that taxes actually *decrease* when an adverse fiscal shock hits. Thus, if we evaluate actual policy from the normative lens of an economy that generates a higher market price of risk, the following conclusion emerges: actual fiscal policy is even worse than we thought.

The basic insights of optimal fiscal policy with recursive utility hold also in an economy with capital as in the setups of [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#). The planner still over-insures and sets high and persistent labor taxes against good shocks. Furthermore, in contrast to the essentially zero ex-ante capital tax result of [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#), there is an incentive to introduce an ex-ante capital *subsidy*. The reason is simple: the planner again mitigates fiscal shocks and manipulates prices by using essentially a state-contingent subsidy to capital income in bad times and a state-contingent capital tax in good times. Bad times are weighed more though due to high marginal utility and a high marginal product of capital. Thus, the weighted *average* of the state-contingent intertemporal distortions becomes negative, leading to an ex-ante subsidy.

**Related literature.** The main reference on optimal taxation with time-additive expected utility for an economy without capital is [Lucas and Stokey \(1983\)](#). The respective references for an economy with capital are [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#). The models I examine reduce to the models analyzed in these studies, if I equate the risk aversion parameter to the inverse of the intertemporal elasticity of substitution parameter. Furthermore, the economy with capital reduces to the deterministic economy of [Chamley \(1986\)](#), if I shut off uncertainty.<sup>2</sup>

---

<sup>2</sup>It is worth noting that [Chamley](#) demonstrated the generality of the zero capital tax result at the deterministic steady state by using the preferences of [Koopmans \(1960\)](#). See [Chari and Kehoe \(1999\)](#) for a comprehensive survey

Related studies include [Farhi and Werning \(2008\)](#), who analyze the implications of recursive preferences for private information setups and [Karantounias \(2013\)](#), who analyzes optimal taxation in an economy without capital, in a setup where the household entertains fears of misspecification but the fiscal authority does not. Of interest is also the work of [Gottardi et al. \(2015\)](#), who study optimal taxation of human and physical capital with uninsurable idiosyncratic shocks and recursive preferences.<sup>3</sup>

Other studies have analyzed the interaction of fiscal policies and asset prices with recursive preferences from a positive angle. [Gomes et al. \(2013\)](#) build a quantitative model and analyze the implications of fiscal policies on asset prices and the wealth distribution. [Croce et al. \(2012a\)](#) show that corporate taxes can create sizeable risk premia with recursive preferences. [Croce et al. \(2012b\)](#) analyze the effect of exogenous fiscal rules on the endogenous growth rate of the economy. None of these studies though considers *optimal* policy.

Another relevant line of research is the analysis of optimal taxation with time-additive expected utility and restricted asset markets as in [Aiyagari et al. \(2002\)](#), [Farhi \(2010\)](#), [Shin \(2006\)](#), [Sleet and Yeltekin \(2006\)](#), [Bhandari et al. \(2016\)](#) or with time-additive expected utility and private information as in [Sleet \(2004\)](#). In the study of [Aiyagari et al. \(2002\)](#), who provide the foundation of the tax-smoothing results of [Barro \(1979\)](#), the lack of insurance markets causes the planner to allocate distortions in a time-varying and persistent way. However, the lack of markets implies that the planner increases the tax rate when government spending is high. Instead, the opposite happens in the current paper.<sup>4</sup> More generally, with incomplete markets as in [Aiyagari et al. \(2002\)](#), the planner would like to allocate tax distortions in a constant way across states and dates but he *cannot*, whereas with complete markets and recursive preferences he could in principle follow a constant distortion policy, but does not find it optimal to do so.

The paper is organized as follows. Section 2 lays out an economy without capital and section 3 sets up the Ramsey problem and its recursive formulation. Section 4 is devoted to the analysis of the *excess burden* of distortionary taxation, a multiplier that reflects how tax distortions are allocated across states and dates. The implications for labor taxes are derived in section 5. Detailed numerical exercises are provided in section 6. Section 7 analyzes government debt returns and optimal fiscal insurance. Section 8 extends the analysis to an economy with capital and considers the optimal ex-ante capital tax. Section 9 discusses the case of preference for late resolution of uncertainty. Finally, section 10 concludes and an Appendix follows. A separate Online Appendix provides additional details and robustness exercises.

---

of optimal fiscal policy.

<sup>3</sup>There is an extensive literature that studies optimal risk-sharing with recursive utility. See [Anderson \(2005\)](#) and references therein.

<sup>4</sup>Furthermore, with incomplete markets as in [Aiyagari et al. \(2002\)](#), it is typically optimal to front-load distortions in order to create a buffer stock of assets, furnishing a tax rate with a negative drift. In contrast, in the current analysis the tax rate exhibits a positive drift, in order to take advantage of cheaper state-contingent debt. It is interesting to observe that [Sleet \(2004\)](#) also obtains a positive drift in the tax rate in a setup with private information about the government spending needs.

## 2 Economy without capital

I start the analysis of optimal fiscal policy with recursive utility in an economy without capital as in [Lucas and Stokey \(1983\)](#). In a later section, I extend the analysis to an economy with capital as in [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#) and I derive the implications for capital taxation.

Time is discrete and the horizon is infinite. There is uncertainty in the economy stemming from exogenous government expenditure shocks  $g$ . Shocks take values in a finite set. Let  $g^t \equiv (g_0, g_1, \dots, g_t)$  denote the partial history of shocks up to time  $t$  and let  $\pi_t(g^t)$  denote the probability of this history. The initial shock is assumed to be given, so that  $\pi_0(g_0) = 1$ .

The economy is populated by a representative household that is endowed with one unit of time and consumes  $c_t(g^t)$ , works  $h_t(g^t)$ , pays linear labor income taxes with rate  $\tau_t(g^t)$  and trades in complete asset markets. Leisure of the household is  $l_t(g^t) = 1 - h_t(g^t)$ . The notation denotes that the relevant variables are measurable functions of the history  $g^t$ . Labor markets are competitive, which leads to an equilibrium wage of unity,  $w_t(g^t) = 1$ . The resource constraint in the economy reads

$$c_t(g^t) + g_t = h_t(g^t), \forall t, g^t. \quad (1)$$

### 2.1 Preferences

The representative household ranks consumption and leisure plans following a recursive utility criterion of [Kreps and Porteus \(1978\)](#). I focus on the isoelastic preferences of [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#) (EZW henceforth), that are described by the utility recursion

$$V_t = [(1 - \beta)u(c_t, 1 - h_t)^{1-\rho} + \beta(E_t V_{t+1}^{1-\gamma})^{\frac{1-\rho}{1-\gamma}}]^{\frac{1}{1-\rho}}, \quad (2)$$

where  $u(c, 1 - h) > 0$ . The household derives utility from a composite good that consists of consumption and leisure,  $u(c, 1 - h)$ , and from the certainty equivalent of continuation utility,  $\mu_t \equiv (E_t V_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}$ .  $E_t$  denotes the conditional expectation operator given information at  $t$  with respect to measure  $\pi$ . The parameter  $1/\rho$  captures the constant intertemporal elasticity of substitution between the composite good and the certainty equivalent, whereas the parameter  $\gamma$  represents risk aversion with respect to atemporal gambles in continuation values. These preferences reduce to standard time-additive expected utility when  $\rho = \gamma$ . This is easily seen by applying the monotonic transformation  $v_t \equiv \frac{V_t^{1-\rho} - 1}{(1-\beta)(1-\rho)}$ , since the utility recursion (2) becomes

$$v_t = U(c_t, 1 - h_t) + \beta \frac{\left[ E_t \left[ 1 + (1 - \beta)(1 - \rho)v_{t+1} \right]^{\frac{1-\gamma}{1-\rho}} \right]^{\frac{1-\rho}{1-\gamma}} - 1}{(1 - \beta)(1 - \rho)}, \quad (3)$$

where  $U(c, 1-h) \equiv \frac{u^{1-\rho}-1}{1-\rho}$ . Recursion (3) implies that the household is *averse* to volatility in future utility when  $\rho < \gamma$ , whereas it *loves* volatility when  $\rho > \gamma$ .<sup>5</sup> Thus, when  $\rho < \gamma$ , recursive utility adds *curvature* with respect to future risks, a feature that is typically necessary to reproduce asset-pricing facts.<sup>6</sup> For that reason, I assume  $\rho < \gamma$  for the main body of the paper, unless otherwise specified. In a later section I consider also the case of  $\rho > \gamma$ .

When  $\rho = 1$ , recursion (2) becomes  $V_t = u_t^{1-\beta} \mu_t^\beta$ . Using the transformation  $v_t \equiv \frac{\ln V_t}{1-\beta}$  we get

$$v_t = \ln u(c_t, 1-h_t) + \frac{\beta}{(1-\beta)(1-\gamma)} \ln E_t \exp [(1-\beta)(1-\gamma)v_{t+1}], \quad (4)$$

which for  $\gamma > 1$  has the interpretation of a *risk-sensitive* recursion with risk-sensitivity parameter  $\sigma \equiv (1-\beta)(1-\gamma)$ .<sup>7</sup>

It will be useful to define

$$m_{t+1} \equiv \left( \frac{V_{t+1}}{\mu_t} \right)^{1-\gamma} = \frac{V_{t+1}^{1-\gamma}}{E_t V_{t+1}^{1-\gamma}}, t \geq 0, \quad (5)$$

with  $m_0 \equiv 1$ . For  $\rho = 1$ , the corresponding definition is  $m_{t+1} = \frac{\exp[(1-\beta)(1-\gamma)v_{t+1}]}{E_t \exp[(1-\beta)(1-\gamma)v_{t+1}]}$ . Note that  $m_{t+1}$  is positive since  $V_{t+1}$  is positive, and that  $E_t m_{t+1} = 1$ . So  $m_{t+1}$  can be interpreted as a change of measure with respect to the conditional probability  $\pi_{t+1}(g_{t+1}|g^t)$ , or, in other words, a conditional likelihood ratio. Similarly, define the product of the conditional likelihood ratios as  $M_t(g^t) \equiv \prod_{i=1}^t m_i(g^i)$ ,  $M_0 \equiv 1$ . This object is a martingale with respect to  $\pi$ ,  $E_t M_{t+1} = M_t$ , and has the interpretation of an unconditional likelihood ratio,  $EM_t = 1$ . I refer to  $\pi_t \cdot M_t$  as the *continuation-value* adjusted probability measure.

---

<sup>5</sup>Define the monotonic function  $H(x) \equiv \left[ (1 + (1-\beta)(1-\rho)x)^{\frac{1-\gamma}{1-\rho}} - 1 \right] / [(1-\beta)(1-\gamma)]$ . Recursion (3) can be written as  $v_t = U_t + \beta H^{-1}(E_t H(v_{t+1}))$ .  $H(x)$  is concave for  $\rho < \gamma$  and convex for  $\rho > \gamma$ . The aversion or love of utility volatility correspond respectively to preference for early or late resolution of uncertainty. They contrast to the case of  $\rho = \gamma$ , which features neutrality to future risks and therefore indifference to the temporal resolution of uncertainty.

<sup>6</sup>See for example Tallarini (2000), Bansal and Yaron (2004), Piazzesi and Schneider (2007) and Epstein et al. (2014).

<sup>7</sup>More generally, in the case of risk-sensitive preferences, the period utility function is not restricted to be logarithmic and the recursion takes the form  $v_t = U_t + \frac{\beta}{\sigma} \ln E_t \exp(\sigma v_{t+1})$ ,  $\sigma < 0$ . There is an intimate link between the risk-sensitive recursion and the multiplier preferences of Hansen and Sargent (2001) that capture the decision maker's fear of misspecification of the probability model  $\pi$ . See Strzalecki (2011) and Strzalecki (2013) for a decision-theoretic treatment of the multiplier preferences and an analysis of the relationship between ambiguity aversion and temporal resolution of uncertainty respectively.

## 2.2 Competitive equilibrium

**Household's problem.** Let  $\{x\} \equiv \{x_t(g^t)\}_{t \geq 0, g^t}$  stand for the sequence of an arbitrary random variable  $x_t$ . The representative household chooses  $\{c, h, b\}$  to maximize  $V_0(\{c\}, \{h\})$  subject to

$$c_t(g^t) + \sum_{g^{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1}) \leq (1 - \tau_t(g^t)) h_t(g^t) + b_t(g^t), \quad (6)$$

the non-negativity constraint for consumption  $c_t(g^t) \geq 0$  and the feasibility constraint for labor  $h_t(g^t) \in [0, 1]$ , where initial debt  $b_0$  is given. The variable  $b_{t+1}(g^{t+1})$  stands for the holdings at history  $g^t$  of an Arrow claim that delivers one unit of consumption next period if the state is  $g_{t+1}$  and zero units otherwise. This security trades at price  $p_t(g_{t+1}, g^t)$  in units of the history-contingent consumption  $c_t(g^t)$ .

The household is also facing a no-Ponzi-game condition that takes the form

$$\lim_{t \rightarrow \infty} \sum_{g^{t+1}} q_{t+1}(g^{t+1}) b_{t+1}(g^{t+1}) \geq 0 \quad (7)$$

where  $q_t(g^t) \equiv \prod_{i=0}^{t-1} p_i(g_{i+1}, g^i)$  and  $q_0 \equiv 1$ . In other words,  $q_t$  stands for the price of an Arrow-Debreu contract at  $t = 0$ .

**Government.** The government taxes labor income and issues state-contingent debt in order to finance the exogenous government expenditures. The dynamic budget constraint of the government takes the form

$$b_t(g^t) + g_t = \tau_t(g^t) h_t(g^t) + \sum_{g^{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1}).$$

When  $b_t > 0$ , the government borrows from the household and when  $b_t < 0$ , the government lends to the household.

**Definition 1.** *A competitive equilibrium with taxes is a stochastic process for prices  $\{p\}$ , an allocation  $\{c, h, b\}$  and a government policy  $\{g, \tau, b\}$  such that: 1) Given prices  $\{p\}$  and taxes  $\{\tau\}$ , the allocation  $\{c, h, b\}$  solves the household's problem. 2) Prices are such that markets clear, i.e. the resource constraint (1) holds.*

## 2.3 Household's optimality conditions

The labor supply decision of the household is governed by

$$\frac{U_l(g^t)}{U_c(g^t)} = 1 - \tau_t(g^t), \quad (8)$$

which equates the marginal rate of substitution between consumption and leisure with the after-tax wage. The first-order condition with respect to an Arrow security equates its price to the household's intertemporal marginal rate of substitution,

$$\begin{aligned} p_t(g_{t+1}, g^t) &= \beta \pi_{t+1}(g_{t+1}|g^t) \left( \frac{V_{t+1}(g^{t+1})}{\mu_t} \right)^{\rho-\gamma} \frac{U_c(g^{t+1})}{U_c(g^t)} \\ &= \beta \pi_{t+1}(g_{t+1}|g^t) m_{t+1}(g^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_c(g^{t+1})}{U_c(g^t)}, \end{aligned} \quad (9)$$

where the second line uses the definition of the conditional likelihood ratio (5). The transversality condition is

$$\lim_{t \rightarrow \infty} \sum_{g^{t+1}} \beta^{t+1} \pi_{t+1}(g^{t+1}) M_{t+1}(g^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} U_c(g^{t+1}) b_{t+1}(g^{t+1}) = 0. \quad (10)$$

The stochastic discount factor  $S_{t+1}$  with EZW utility is

$$S_{t+1} \equiv \beta \left( \frac{V_{t+1}}{\mu_t} \right)^{\rho-\gamma} \frac{U_{c,t+1}}{U_{ct}} = \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_{c,t+1}}{U_{ct}}. \quad (11)$$

The stochastic discount factor features *continuation values*, scaled by their certainty equivalent  $\mu_t$ , when  $\rho \neq \gamma$ . Besides caring for the short-run ( $U_{c,t+1}/U_{ct}$ ), the household cares also for the “long-run,” in the sense that the *entire* sequence of future consumption and leisure – captured by continuation values – *directly* affects  $S_{t+1}$ . Increases in consumption growth at  $t + 1$  reduce period marginal utility and therefore the stochastic discount factor in the standard time-additive setup. When  $\rho < \gamma$ , increases in continuation values act exactly the *same* way; they decrease the stochastic discount factor, because the household dislikes volatility in future utility. This is the essence of the additional “curvature” that emerges with recursive utility.<sup>8</sup>

### 3 Ramsey problem

The Ramsey planner maximizes at  $t = 0$  the utility of the representative household over the set of competitive equilibrium allocations. Competitive equilibrium allocations are characterized by resource constraints, budget constraints and optimality conditions that involve equilibrium prices and taxes. I follow the primal approach of [Lucas and Stokey \(1983\)](#) and use the optimality conditions to replace after-tax wages and prices with the respective marginal rates of substitution. As a result, I formulate a policy problem where the planner chooses allocations that satisfy the resource constraint (1) and *implementability* constraints, i.e. constraints that allow the allocation

---

<sup>8</sup>[Bansal and Yaron \(2004\)](#) and [Hansen et al. \(2008\)](#) have explored ways of making the continuation value channel quantitatively important in order to increase the market price of risk.

to be implemented as a competitive equilibrium.

### 3.1 Implementability constraints

The household’s dynamic budget constraint (6) holds with equality. Use (8) and (9) to eliminate labor taxes and equilibrium prices from the constraint to get a sequence of implementability constraints:

**Proposition 1.** *The Ramsey planner faces the following implementability constraints:*

$$U_{ct}b_t = U_{ct}c_t - U_{lt}h_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}b_{t+1}, t \geq 0$$

where  $c_t \geq 0$ ,  $h_t \in [0, 1]$  and  $(b_0, g_0)$  given. Furthermore, the transversality condition (10) has to be satisfied. The conditional likelihood ratios  $m_{t+1} = V_{t+1}^{1-\gamma}/E_t V_{t+1}^{1-\gamma}$ ,  $t \geq 0$ , are determined by continuation values that follow recursion (2).

Complete markets allow the collapse of the household’s dynamic budget constraint to a unique intertemporal budget constraint. However, maintaining the dynamic budget constraint is convenient for a recursive formulation, as we will see in the next section.

**Definition 2.** *The Ramsey problem is to maximize at  $t = 0$  the utility of the representative household subject to the implementability constraints of proposition 1 and the resource constraint (1).*

### 3.2 Recursive formulation

I follow the methodology of Kydland and Prescott (1980) and break the Ramsey problem in two subproblems: the problem from period one onward and the initial period problem. Let  $z_t$  denote debt in (period) marginal utility units,  $z_t \equiv U_{ct}b_t$ .<sup>9</sup> I represent the policy problem for  $t \geq 1$  recursively by keeping track of  $g$  – the exogenous shock– and  $z$ , the variable that captures the commitment of the planner to his past promises. Note that  $z$  is a forward-looking variable that is not inherited from the past. This creates the need to specify  $Z(g)$ , the space where  $z$  lives. The set  $Z(g)$  represents the values of debt in marginal utility units that can be generated from an implementable allocation when the shock is  $g$  and is defined in the Appendix.<sup>10</sup> Let  $V(z_1, g_1)$  denote the value function of the planner’s problem from period one onward, where  $z_1 \in Z(g_1)$  and assume that shocks follow a Markov process with transition probabilities  $\pi(g'|g)$ .

<sup>9</sup>In the interest of brevity, I sometimes skip the “marginal utility units” qualification and refer to  $z$  simply as debt. The meaning is always clear from the context.

<sup>10</sup>A separate Online Appendix provides the sequential formulation of the Ramsey problem.

**Bellman equation.** The functional equation that determines the value function  $V$  takes the form

$$V(z, g) = \max_{c, h, z'_{g'}} \left[ (1 - \beta)u(c, 1 - h)^{1-\rho} + \beta \left[ \sum_{g'} \pi(g'|g) V(z'_{g'}, g')^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

subject to

$$z = U_c c - U_l h + \beta \sum_{g'} \pi(g'|g) \frac{V(z'_{g'}, g')^{\rho-\gamma}}{\left[ \sum_{g'} \pi(g'|g) V(z'_{g'}, g')^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z'_{g'} \quad (12)$$

$$c + g = h \quad (13)$$

$$c \geq 0, h \in [0, 1] \quad (14)$$

$$z'_{g'} \in Z(g'). \quad (15)$$

The planner is maximizing welfare by choosing consumption, labor (and thus effectively the labor tax), and next period's state-contingent debt in marginal utility units,  $z'_{g'}$ , subject to the government budget constraint (12) (expressed in terms of allocations), and the resource constraint, (13). The nature of the Ramsey problem is fundamentally changed because, in contrast to time-additive utility, continuation values matter for the determination of the market value of the government debt portfolio, and therefore show up in constraint (12). As such, the dynamic tradeoff of taxing at the current period versus postponing taxation by issuing debt is *altered*, since the planner has now to take into account how new debt issuance affects equilibrium prices through the “long-run.” This tradeoff is at the heart of next section.

**Initial period problem.** The value of  $z_1$  that was taken as given in the formulation of the planner's problem at  $t \geq 1$  is chosen optimally at  $t = 0$ . In this sense,  $z$  is a *pseudo-state* variable, i.e. a jump variable that is treated as a state variable in order to capture the commitment of the planner to the optimal plan devised at the initial period. The initial period problem is stated in the Online Appendix.

## 4 Recursive utility and the excess burden of taxation

### 4.1 Overview of the mechanism

How does the government tax across states and dates and how does it manage its state-contingent debt in a welfare-maximizing way? To fix ideas, I provide here an overview of the mechanism that is supported by the analysis of the optimality conditions and the numerical analysis of later sections.

The government is absorbing spending shocks through its debt portfolio. It achieves that by selling claims to consumption against low spending shocks (good times) and by purchasing claims to consumption against high spending shocks (bad times). In the standard time-additive setup, the size of sales and purchases of state-contingent claims is such that the tax rate remains essentially constant across states and dates, leading to the typical tax-smoothing result. Note that consumption is high (low), and therefore the stochastic discount factor is low (high) when spending shocks are low (high). So the price of claims sold is low and the price of claims bought is high.

The government has similar motives to use state-contingent debt in order to hedge fiscal risks in a recursive utility economy. The difference is that there is a novel instrument to affect the stochastic discount factor, lifetime utilities, which allows the government to make debt cheaper, *amplifying* fiscal hedging: the government “over-insures” by selling more claims to consumption against low shocks relative to the time-additive benchmark. Issuance of more debt against low spending shocks reduces continuation utilities and, therefore, *increases* the stochastic discount factor more than in the time-additive case, increasing the price of claims sold. Consequently, the current revenue from selling claims to the private sector against a low spending shock next period increases, allowing the relaxation of the government budget and less taxation at the current period. More claims sold against a low shock next period implies that higher taxes have to be levied in the future at that state, in order to repay debt. A similar mechanism holds for high spending states: the government insures against fiscal risk by purchasing more claims to consumption against high spending shocks relative to the time-additive economy. These actions increase the household’s utility, depressing therefore the stochastic discount factor and the price of claims bought. More assets (or less debt) against high shocks implies less taxes contingent on these states of the world.

The mechanism is intuitive and makes economic sense. It simply says that the planner should mitigate the effects of fiscal shocks by taxing more in good times and less in bad times. By doing that, state-contingent debt against good times becomes cheaper and state-contingent assets against bad times become more profitable, due to the additional curvature of recursive utility. Furthermore, this mechanism leads on average to back-loading of tax distortions over time, due to the reduced interest rate cost of debt. Lastly, persistence of optimal tax rates is optimal independent of the persistence of exogenous shocks: the planner changes smoothly the tax rate over time in order to take full advantage of the forward-looking nature of continuation utilities.

## 4.2 Preliminaries: expected utility and the excess burden

Consider now the specifics of the mechanism. For the analysis of the problem it is easier to use the transformed value function,  $v(z, g) \equiv \frac{V(z, g)^{1-\rho} - 1}{(1-\beta)(1-\rho)}$ , that corresponds to recursion (3). The *entire* action is coming from  $\Phi$ , the multiplier on the implementability constraint of the transformed problem.

The envelope condition is  $v_z(z, g) = -\Phi \leq 0$ , since  $\Phi$  is non-negative,  $\Phi \geq 0$ .<sup>11</sup> So  $\Phi$  captures the *cost* of an additional unit of debt in marginal utility units. Increases in debt are costly because they have to be accompanied with an increase in distortionary taxation ( $\Phi = 0$  when lump-sum taxes are available).<sup>12</sup>

I refer to  $\Phi$  as the *excess burden* of distortionary taxation and interpret it as an indicator of tax *distortions*. In order to build intuition about its role, consider first the time-additive expected utility world of Lucas and Stokey (1983) where  $\rho = \gamma$ . The optimality condition with respect to new debt  $z'_{g'}$  takes the form

$$-v_z(z'_{g'}, g') = \Phi. \quad (16)$$

Optimality condition (16) has a typical marginal cost and marginal benefit interpretation. The left-hand side captures the marginal cost of issuing more debt against  $g'$  next period. Selling more claims to consumption at  $g'$  is costly because the planner has to increase distortionary taxation in order to repay debt. However, by issuing more debt for next period, the planner can relax the government budget and tax less at the current period. The marginal benefit of relaxing the budget constraint has shadow value  $\Phi$ , which is the right-hand side of (16).

By using the envelope condition, condition (16) implies that  $\Phi'_{g'} = \Phi, \forall g'$ , for all values of the state  $(z, g)$ . Thus, in a time-additive expected utility economy, the planner sells and buys as many state-contingent claims as necessary, in order to *equalize* the excess burden of taxation across states and dates. This is the formal result that hides behind the *tax-smoothing* intuition in typical frictionless Ramsey models. Furthermore, the constant excess burden is also the source of Lucas and Stokey's celebrated *history-independence* result, since optimal allocations and tax rates can be written solely as functions of the exogenous shocks and the constant  $\Phi$ .

### 4.3 Pricing with recursive utility and the excess burden

Turn now to the recursive utility case. New debt issuance at  $g'$  is governed by the following optimality condition:

$$\underbrace{-v_z(z'_{g'}, g')}_{\text{MC of increasing } z'_{g'}} = \Phi \cdot \left[ \underbrace{1}_{\text{EU term}} + \underbrace{(1 - \beta)(\rho - \gamma)v_z(z'_{g'}, g')\eta'_{g'}}_{\text{EZW term: price effect of increasing } z'_{g'}} \right], \quad (17)$$

<sup>11</sup>I am implicitly assuming that the government has access to lump-sum transfers, so that the dynamic implementability constraint takes the form  $z_t \leq U_{ct}c_t - U_{lt}h_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$ .

<sup>12</sup> $\Phi$  would also be zero if the government had sufficient initial assets that could support the first-best allocation. This case is ruled out here in order to have an interesting second-best problem.

where

$$\eta'_{g'} \equiv V'_{g'}{}^{\rho-1} z'_{g'} - \mu^{\rho-1} \sum_{g'} \pi(g'|g) m'_{g'}{}^{\frac{\rho-\gamma}{1-\gamma}} z'_{g'}. \quad (18)$$

Equivalently, by using the definition of  $m'_{g'}$ , we can rewrite the variable  $\eta'_{g'}$  as  $\eta'_{g'} = V'_{g'}{}^{\rho-1} z'_{g'} - \sum_{g'} \pi(g'|g) m'_{g'} V'_{g'}{}^{\rho-1} z'_{g'}$ .<sup>13</sup> So  $\eta'_{g'}$  stands for the conditional innovation of  $V'_{g'}{}^{\rho-1} z'_{g'}$  under  $\pi_t \cdot M_t$  and takes positive and negative values with an average of *zero*,  $\sum_{g'} \pi(g'|g) m'_{g'} \eta'_{g'} = 0$ .

For  $\rho = 1$ ,  $\eta'_{g'}$  simplifies to the state-contingent debt position in marginal utility units relative to the value of the government debt portfolio,  $\eta'_{g'} = z'_{g'} - \sum_{g'} \pi(g'|g) m'_{g'} z'_{g'}$ . For that reason, I call  $\eta'_{g'}$  the government's *relative debt position* in marginal utility units.

**Interpretation.** As in the time-additive case, the left-hand side of (17) denotes the marginal cost of issuing more debt against  $g'$  next period. The right-hand side of (17) measures the *utility* benefit (captured by the multiplication with the current multiplier  $\Phi$ ) coming from the government's *marginal revenue* from debt issuance (the expression inside the brackets). The first term in the brackets captures the same direct increase in revenue as in the time-additive setup, coming from selling more debt. The second term is *novel* and is coming from the change in *prices* due to the increased debt position: an increase in debt reduces utility, which *increases* the stochastic discount factor,  $(\rho - \gamma)v_z > 0$  for  $\rho < \gamma$ . This increase in prices, which is multiplied with  $\eta'_{g'}$ , was absent in the time-additive setup, since the “long-run” was not priced.<sup>14</sup>

How the planner is going to use this novel price effect of recursive utility depends on the *relative debt position*  $\eta'_{g'}$ , according to (17). To see clearly the mechanism, turn into sequence notation, collect the terms that involve  $v_z$ , and use the envelope condition in order to rewrite (17) in terms of the *inverse* of  $\Phi$  (assuming that  $\Phi$  is not zero),<sup>15</sup>

$$\frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} + (1 - \beta)(\rho - \gamma)\eta_{t+1}, t \geq 0, \quad (19)$$

where  $\eta_{t+1} \equiv V_{t+1}{}^{\rho-1} z_{t+1} - \mu_t^{\rho-1} E_t m_{t+1}{}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1} = V_{t+1}{}^{\rho-1} z_{t+1} - E_t m_{t+1} V_{t+1}{}^{\rho-1} z_{t+1}$ . Consider fiscal shocks  $\hat{g}$  and  $\tilde{g}$  at  $t + 1$  such that  $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$ . Then, (19) implies that  $\Phi_{t+1}(\hat{g}) > \Phi_t > \Phi_{t+1}(\tilde{g})$

<sup>13</sup>In definition (18) recall that  $m'_{g'}$  stands for the conditional likelihood ratio,  $\mu$  for the certainty equivalent and  $V'_{g'}$  is shorthand for  $V(z'_{g'}, g')$ . I use the non-transformed value function  $V$  in (18) (which is equal to  $[1 + (1 - \beta)(1 - \rho)v]^{\frac{1}{1-\rho}}$ ) as a matter of convenience; it allows a more compact exposition of the first-order conditions.

<sup>14</sup>The second term in the brackets of the right-hand side of (17) would be absent also in a deterministic economy, since  $\eta'_{g'} \equiv 0, \forall g'$  in that case. This would imply again a constant excess burden of taxation. Thus, apart from the *level* of the constant  $\Phi$ , there is no essential difference between a deterministic world and a stochastic but time-additive world with  $\rho = \gamma$ .

<sup>15</sup>Otherwise, write the optimality condition as  $\Phi_{t+1} = \Phi_t / [1 + (1 - \beta)(\rho - \gamma)\eta_{t+1}\Phi_t]$ . Thus, if  $\Phi_t = 0$ , then  $\Phi_{t+i} = 0, i \geq 0$ , so the first-best is an absorbing state.

for  $\rho < \gamma$ . So, in contrast to the time-additive setup, the excess burden of taxation, and therefore the tax rate, varies across states and dates and is higher at states of the world next period, against which the relative debt position is positive, and lower at states of the world, against which the relative debt position is negative.<sup>16</sup>

What is happening here? Exactly the story that we highlighted in the overview of the mechanism. The increase in prices due to the additional curvature of recursive utility is *beneficial* at states of the world against which the planner issues relatively more debt. In other words, the planner should *optimally* increase taxes at states of the world next period, against which it is *cheaper* today to issue debt. The opposite happens for states of the world against which the relative debt position is small.

Two comments are due. First, note that it is not just the debt position (adjusted by marginal utility – and continuation utility when  $\rho \neq 1$ ) but the debt position *relative* to (a multiple of) the market value of the debt portfolio,  $E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$ , that matters for the increase or decrease of the excess burden of taxation across states and dates. The reason for this is coming from the state *non-separabilities* that emerge with recursive utility. In particular, an increase of  $z'_{g'}$  may increase the price of the respective claim at  $g'$  by reducing utility, but reduces also the certainty equivalent and *decreases* therefore the rest of the prices of state-contingent claims at  $\bar{g}, \bar{g} \neq g'$ . This is why the *relative* position  $\eta_{t+1}$  captures the *net* effect of price manipulation through the continuation utility channel.

Second, in the overview of the mechanism we stressed that the government is using state-contingent debt to hedge fiscal shocks by selling claims against low spending shocks and purchasing claims (or selling less claims) against high spending shocks. Thus, we expect to have  $b_{t+1}(g_L) > b_{t+1}(g_H)$  for  $g_H > g_L$ . Assume that  $\rho = 1 < \gamma$  and that the same ranking of debt positions holds also for debt in marginal utility units, i.e.  $z_{t+1}(g_L) > z_{t+1}(g_H)$ . Then,  $\eta_{t+1}(g_L) > 0 > \eta_{t+1}(g_H)$ , which implies that  $\Phi_{t+1}(g_L) > \Phi_t > \Phi_{t+1}(g_H)$ . Consequently, the excess burden, and therefore the tax rate, increases for low fiscal shocks and decreases for high fiscal shocks, leading to larger surpluses and deficits. We are going to see explicitly this fiscal hedging when we solve the model numerically.

To conclude this section, the following proposition summarizes the results about the excess burden of taxation.

**Proposition 2.** 1. *The excess burden is constant across states and dates when  $\rho = \gamma$ .*

2. *Assume  $\rho < \gamma$  and let  $\hat{g}$  and  $\tilde{g}$  be shocks at  $t + 1$  such that  $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$ . Then, the law of motion of the excess burden (19) implies that  $\Phi_{t+1}(\hat{g}) > \Phi_t > \Phi_{t+1}(\tilde{g})$ .*

---

<sup>16</sup>The varying excess burden has also implications for the size of  $z_t$  over time. It is tempting to deduce that the planner is not only increasing the excess burden for a high-debt state next period ( $\eta_{t+1} > 0$ ), but also issues more state-contingent debt for next period. Formally, the deduction would be  $\Phi_{t+1} = -v_z(z_{t+1}, g) > \Phi_t = -v_z(z_t, g) \Rightarrow z_{t+1} > z_t$ , which is a statement about the concavity of  $v$  at  $g$ . This statement cannot be made in general due to the non-convexities of the Ramsey problem, but it turns out to be numerically true.

3. (Fiscal hedging and the excess burden) Let  $g_H > g_L$  and assume that  $\rho = 1 < \gamma$ . If  $z_{t+1}(g_L) > z_{t+1}(g_H)$ , then  $\Phi_{t+1}(g_L) > \Phi_t > \Phi_{t+1}(g_H)$ .<sup>17</sup>

#### 4.4 Dynamics of the excess burden of taxation

The relative debt position  $\eta_t$  captures the incentives of the planner to increase or decrease the excess burden, *given* the past shadow cost of debt and tax promises,  $\Phi_{t-1}$ . This fact introduces dependence on the history of shocks. To understand the role of the past, consider a change in debt at time  $t$ . This change will affect continuation values at  $t$  but also at all previous periods, because utilities are forward-looking: the household at  $t - i, i = 1, 2, \dots, t$  is taking into account the entire future stream of consumption and leisure when it prices Arrow claims. As a result, *all* past prices of state-contingent claims  $p_i(s_{i+1}, s^i), i = 0, 1, 2, \dots, t - 1$  change with a change in continuation values at time  $t$ . This is why the excess burden depends on the *sum* of the past relative debt positions  $\{\eta_i\}_{i=1}^t$ , which can be seen by solving (19) backwards. Furthermore, we have:

**Proposition 3.** (Persistence and back-loading of the excess burden)

The inverse of  $\Phi_t$  is a martingale with respect to the continuation-value adjusted measure  $\pi_t \cdot M_t$  for  $\rho \leq \gamma$ . Therefore,  $\Phi_t$  is a submartingale with respect to  $\pi_t \cdot M_t$ ,  $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$ . As a result,

$$E_t \Phi_{t+1} \geq \Phi_t - \text{Cov}_t(m_{t+1}, \Phi_{t+1}), \quad (20)$$

so if  $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) \leq 0$ ,  $\Phi_t$  is a submartingale with respect to  $\pi$ ,  $E_t \Phi_{t+1} \geq \Phi_t$ .

*Proof.* Take conditional expectation in (19) to get

$$E_t m_{t+1} \frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} E_t m_{t+1} + (1 - \beta)(\rho - \gamma) E_t m_{t+1} \eta_{t+1} = \frac{1}{\Phi_t},$$

since  $E_t m_{t+1} = 1$  and  $E_t m_{t+1} \eta_{t+1} = 0$ . Thus  $1/\Phi_t$  is a martingale with respect to  $\pi_t \cdot M_t$ . Furthermore, since the function  $f(x) = 1/x$  is convex for  $x > 0$ , an application of the conditional version of Jensen's inequality leads to  $E_t m_{t+1} \frac{1}{\Phi_{t+1}} \geq \frac{1}{E_t m_{t+1} \Phi_{t+1}}$ . Set now  $x_t = 1/\Phi_t$  and use the martingale result to finally get  $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$ . Inequality (20) is derived by using the submartingale result and the fact that  $E_t m_{t+1} \Phi_{t+1} = \text{Cov}_t(m_{t+1}, \Phi_{t+1}) + E_t \Phi_{t+1}$ , since  $E_t m_{t+1} = 1$ .  $\square$

The martingale result about the inverse of the excess burden of taxation implies *persistence* independent of the stochastic properties of exogenous shocks, in contrast to the standard time-

<sup>17</sup>Clearly, the corresponding statement for  $\rho \neq 1 < \gamma$  is: if  $V_{t+1}^{\rho-1}(g_L) z_{t+1}(g_L) > V_{t+1}^{\rho-1}(g_H) z_{t+1}(g_H)$ , then  $\Phi_{t+1}(g_L) > \Phi_t > \Phi_{t+1}(g_H)$ .

additive Ramsey results.<sup>18</sup> Furthermore, the submartingale result shows that the planner wants on “average” to *back-load* tax distortions, in the sense that the excess burden exhibits a positive drift with respect to the continuation-value adjusted measure, independent of  $\rho \leq \gamma$ . In order to determine the drift with respect to the actual measure that generates uncertainty,  $\pi$ , we need to determine the covariance of the excess burden with the change of measure  $m_{t+1}$ . Consider without loss of generality the case of  $\rho = 1 < \gamma$ . Then, high fiscal shocks, since they reduce utility, are associated with a higher conditional probability mass and therefore a higher  $m_{t+1}$ , leading to a positive correlation of  $m_{t+1}$  with spending. Furthermore, we expect the excess burden to be negatively correlated with spending. As a result, we expect  $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) \leq 0$  and therefore proposition 3 implies a positive drift in  $\Phi_t$  with respect to  $\pi$ . More intuitively, since the average excess burden of taxation is increasing according to the utility-adjusted beliefs that do *not* assign a lot of probability mass on states of the world with a high excess burden, it will still be increasing on average according to the data-generating process, which puts more weight on exactly these contingencies of a high excess burden. We will explore further the persistence and the back-loading of tax distortions in the numerical exercises section.

## 5 Optimal labor income taxation

The following proposition exhibits the exact relationship of the excess burden of taxation with the labor tax.

**Proposition 4.** (*Labor tax*) *The optimal labor tax is*

$$\tau_t = \Phi_t \frac{\epsilon_{cc,t} + \epsilon_{ch,t} + \epsilon_{hh,t} + \epsilon_{hc,t}}{1 + \Phi_t(1 + \epsilon_{hh,t} + \epsilon_{hc,t})}, \quad t \geq 1.$$

where  $\epsilon_{cc} \equiv -U_{cc}/U_c > 0$  and  $\epsilon_{ch} \equiv U_{cl}h/U_c$ , i.e. the own and cross elasticity of the period marginal utility of consumption, and  $\epsilon_{hh} \equiv -U_{ll}h/U_l > 0$  and  $\epsilon_{hc} \equiv U_{lc}c/U_l$ , the own and cross elasticity of the period marginal disutility of labor. When  $U_{cl} \geq 0$ , then  $\epsilon_{ch}, \epsilon_{hc} \geq 0$  and  $\tau_t \geq 0$ .<sup>19</sup>

*Proof.* Let  $\Omega(c, h) \equiv U_c(c, 1 - h)c - U_l(c, 1 - h)h$  stand for consumption net of after-tax labor income, in marginal utility units. This object is in equilibrium equal to the *primary surplus* in marginal utility units. Let  $\lambda$  denote the multiplier on the resource constraint of the Ramsey problem with the transformed value function  $v$ . The first-order necessary conditions with respect to  $(c, h)$  are

<sup>18</sup>In the Online Appendix I discuss why the martingale property is not sufficient to establish convergence results of the inverse of the excess burden with respect to  $\pi$ .

<sup>19</sup>The labor tax formula holds also for the deterministic and stochastic time-additive case for any period utility  $U$  that satisfies the standard monotonicity and concavity assumptions, i.e. without being restricted to  $U = (u^{1-\rho} - 1)/(1 - \rho)$ ,  $u > 0$ .

$$c : U_c + \Phi\Omega_c = \lambda \quad (21)$$

$$h : U_l - \Phi\Omega_h = \lambda, \quad (22)$$

where  $\Omega_i, i = c, h$  denotes the respective partial derivative. Combine the first-order conditions (21)-(22) to get the optimal wedge in labor supply,  $\frac{U_l}{U_c} \cdot \frac{1 - \Phi \frac{\Omega_h}{U_l}}{1 + \Phi \frac{\Omega_c}{U_c}} = 1$ . Associate the derivatives  $\Omega_i, i = c, h$  to elasticities as  $\Omega_c/U_c = 1 - \epsilon_{cc} - \epsilon_{ch}$  and  $\Omega_h/U_l = -1 - \epsilon_{hh} - \epsilon_{hc}$ . Use the labor supply condition  $U_l/U_c = 1 - \tau$  and rewrite the optimal wedge as  $\tau = -\Phi(\Omega_c/U_c + \Omega_h/U_l)/(1 - \Phi\Omega_h/U_l)$ . The result follows.  $\square$

The formula in proposition 4 expresses the optimal labor tax in terms of the excess burden of taxation  $\Phi_t$  and the elasticities of the period marginal utility of consumption and disutility of labor. Ceteris paribus, the labor tax varies *monotonically* with the excess burden of taxation, a fact which justifies the interpretation of  $\Phi_t$  as an indicator of tax distortions.<sup>20</sup> Period elasticities in the optimal tax formula reflect the sensitivity of the surplus in marginal utility units to shocks. They capture the sensitivity of labor supply to changes in the tax rate and the pricing effects of the period marginal utility channel in the stochastic discount factor – the only pricing effect in the time-additive case. Assume for example that  $U_{cl} = 0$ . The optimal tax formula shows that the larger  $\epsilon_{cc}$ , the larger the tax rate, ceteris paribus. The reason is simple. A large tax rate reduces consumption and increases marginal utility, *increasing* therefore the discounted value of surpluses and relaxing the government budget. This is essentially the only type of interest rate manipulation with time-additive utility.<sup>21</sup>

When  $\rho = \gamma$ , we have  $\Phi_t = \Phi$ , and the labor tax varies only due to variation in period elasticities. Thus, when elasticities are *constant*, optimal policy prescribes *perfect* tax-smoothing. With recursive utility though, even in the constant period elasticity case, the labor tax varies monotonically with the non-constant excess burden of taxation. Consider for example the composite good  $u$

$$u(c, 1 - h) = \left[ c^{1-\rho} - (1 - \rho)a_h \frac{h^{1+\phi_h}}{1 + \phi_h} \right]^{\frac{1}{1-\rho}}, \quad (23)$$

<sup>20</sup>We have  $\frac{\partial \tau}{\partial \Phi} |_{\epsilon_{i,j} \text{ constant}} = \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc}}{[1 + \Phi(1 + \epsilon_{hh} + \epsilon_{hc})]^2} > 0$ , as long as the numerator is positive.  $U_{cl} \geq 0$  is sufficient for that.

<sup>21</sup>Similarly, (22) implies that a reduction in labor (through an increase in the tax rate) is –by increasing tax revenues– beneficial when  $\epsilon_{hh}$  is high, i.e. when the Frisch elasticity of labor supply is small. Thus, the labor tax formula in proposition 4 contains the standard static Ramsey prescription of taxing more labor when it is supplied inelastically.

which implies a period utility function with constant elasticities,  $U = \frac{c^{1-\rho}-1}{1-\rho} - a_h \frac{h^{1+\phi_h}}{1+\phi_h}$ .<sup>22</sup> We get the following proposition:

**Proposition 5.** (*Labor tax with power utility and constant Frisch elasticity*)

1. *The labor tax follows the law of motion*

$$\frac{1}{\tau_{t+1}} = \frac{1}{\tau_t} + \frac{(1-\beta)(\rho-\gamma)}{\rho+\phi_h} \eta_{t+1}, t \geq 1. \quad (24)$$

2. *Tax rates across states and dates:*

- *Let  $\rho < \gamma$ . Let  $\hat{g}$  and  $\tilde{g}$  be shocks at  $t+1$  such that  $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$ . Then,  $\tau_{t+1}(\hat{g}) > \tau_t > \tau_{t+1}(\tilde{g})$ .*
- *Let  $\rho = 1 < \gamma$  and assume that shocks take two values,  $g_H > g_L$ . If  $z_{t+1}(g_L) > z_{t+1}(g_H)$ , then  $\tau_{t+1}(g_L) > \tau_t > \tau_{t+1}(g_H)$ .*<sup>23</sup>

3. (*Persistence and back-loading of the labor tax*) *The inverse of the labor tax is a martingale with respect to  $\pi_t \cdot M_t$  for  $\rho \leq \gamma$ . Therefore,  $\tau_t$  is a submartingale with respect to  $\pi_t \cdot M_t$ ,  $E_t m_{t+1} \tau_{t+1} \geq \tau_t$  and*

$$E_t \tau_{t+1} \geq \tau_t - Cov_t(m_{t+1}, \tau_{t+1}).$$

*If  $Cov_t(m_{t+1}, \tau_{t+1}) \leq 0$ , then  $E_t \tau_{t+1} \geq \tau_t$ .*

*Proof.* The labor tax formula in proposition 4 specializes to

$$\tau_t = \frac{\Phi_t(\rho + \phi_h)}{1 + \Phi_t(1 + \phi_h)}, t \geq 1. \quad (25)$$

The formula shows that the crucial parameter for the period elasticities channel is  $\rho$  (and *not*  $\gamma$ ), whereas both  $\rho$  and  $\gamma$  affect the Ramsey outcome through the law of motion of  $\Phi_t$ , (19). Taking inverses in (25) delivers  $\frac{1}{\tau_t} = \frac{1+\phi_h}{\rho+\phi_h} + \frac{1}{\rho+\phi_h} \frac{1}{\Phi_t}$ , so  $1/\tau_t$  is an *affine* function of  $1/\Phi_t$ . Use then (19) to get the law of motion of the labor tax (24). Notice the resemblance of (24) to (19), a fact that leads to the same conclusions about the variation of tax rates across states and dates and (sub)martingale properties as in proposition 3.  $\square$

When we have a period utility function with a power subutility of consumption and constant Frisch elasticity, period elasticities are constant and the labor tax behaves *exactly* as the excess

<sup>22</sup>It is assumed that parameters are such so that  $c^{1-\rho} - (1-\rho)a_h \frac{h^{1+\phi_h}}{1+\phi_h} > 0$ , so that  $u > 0$  is well defined. For  $\rho = 1$ , the utility recursion becomes  $V_t = \exp \left[ (1-\beta)(\ln c - a_h \frac{h^{1+\phi_h}}{1+\phi_h}) + \beta \ln \mu_t \right]$ . If we want to drop these restrictions on preference parameters, we can just consider risk-sensitive preferences with the particular period utility  $U$ .

<sup>23</sup>Same comment applies as in footnote 17.

burden of taxation, following the elegant law of motion (24). The entire analysis of section 4 about the variation of the excess burden across states and dates, the positive drift and persistence, can be recast word by word in terms of the labor tax and will not be repeated.

## 6 Numerical exercises

In this section I provide various numerical exercises in order to highlight three main results of the paper: a) the planner’s “over-insurance” that leads to higher tax rates when fiscal shocks are favorable and smaller tax rates when fiscal shocks are adverse, b) the volatility and back-loading of tax distortions, c) the persistence of tax distortions independent of the persistence of exogenous shocks. In a nutshell, the tax rate behaves like a random walk with a positive drift in the short and medium-run, with an increment that is negatively correlated with fiscal shocks.

### 6.1 Solution method

The numerical analysis with recursive preferences is highly non-trivial. There are three complications: At first, the state space where  $z$  lives is endogenous, i.e. we have to find values of debt in marginal utility units that can be generated at a competitive equilibrium. Second, the contraction property is impaired due to the presence of the value functions in the implementability constraint, a fact which makes convergence of iterative procedures difficult. Third, there are novel non-convexities in the implementability constraint due to recursive utility. I illustrate here the gist of the numerical method and provide additional details in the Online Appendix.

The way I proceed is as follows. I generate feasible values of  $z$  and calculate the respective utility by assuming that the planner follows a constant- $\Phi$  policy, i.e. I assume that the planner ignores the prescriptions of optimal policy and just equalizes the excess burden of taxation over states and dates. By varying  $\Phi$ , I can generate a *set* of values of  $z$ , which I use as a proxy of the state space. The respective value functions are used as a first guess in the numerical algorithm. I implement a double loop: In the inner loop, I fix the value function in the constraint and solve the Bellman equation using grid search. The inner loop is convergent. In the outer loop, I update the value function in the constraint and repeat the inner loop. Although there is no guarantee of convergence of the double loop, this procedure works fairly well. After convergence, I add a final step to improve precision: I employ the output of the double loop as a first guess, fit the value functions with cubic splines and use a continuous optimization routine.

### 6.2 Calibration

I use the utility function of proposition 5 that delivers perfect tax-smoothing in the time-additive economy and a standard calibration. In particular, let  $\rho = 1$  and consider the utility recursion

$$v_t = \ln c_t - a_h \frac{h_t^{1+\phi_h}}{1+\phi_h} + \frac{\beta}{(1-\beta)(1-\gamma)} \ln E_t \exp((1-\beta)(1-\gamma)v_{t+1}), \quad (26)$$

where  $\gamma > 1$ . The frequency is annual and Frisch elasticity is unitary,  $(\beta, \phi_h) = (0.96, 1)$ . The atemporal risk aversion is  $\gamma = 10$ .<sup>24</sup> I assume that shocks are i.i.d. in order to focus on the persistence generated *endogenously* by optimal policy. Expenditures shocks take two values,  $g_L = 0.072$  and  $g_H = 0.088$ , with probability  $\pi = 0.5$ . These values correspond to 18% and 22% of average first-best output respectively, or 20.37% and 24.28% of output in the second-best expected utility economy. So the standard deviation of the share of government spending in output is small and about 2%. I set  $a_h = 7.8125$  which implies that the household works on average 40% of its available time if we are at the first-best, or 35.8% of its time in the second-best, time-additive economy. Initial debt is zero and the initial realization of the government expenditure shock is low,  $g_0 = g_L$ .

### 6.3 Expected utility plan

The time-additive expected utility case of  $\gamma = 1$  corresponds to the environment of [Lucas and Stokey \(1983\)](#). The Ramsey plan is history-independent and the tax rate is constant and equal to 22.3%. The planner issues zero debt against low shocks,  $b_L = 0$ , and insures against high spending by buying *assets*,  $b_H < 0$ . The level of assets corresponds to 3.81% of output. Thus, the debt-to-output ratio has mean  $-1.91\%$  and standard deviation 1.91%. Whenever there is a low shock, the planner, who has no debt to repay ( $b_L = 0$ ), runs a surplus  $\tau h_L - g_L > 0$  and uses the surplus to buy assets against the high shock. The amount of assets is equal to  $b_H = (\tau h_H - g_H)/(1 - \beta\pi)$ . When the shock is high, the planner uses the interest income on these assets to finance the deficit  $\tau h_H - g_H < 0$ .<sup>25</sup>

### 6.4 Fiscal hedging, over-insurance and price manipulation

Turning to recursive utility, the left panel in [figure 1](#) plots the difference between the policy functions for  $z'$  next period when  $g'$  is low and high respectively. The graph shows that the

---

<sup>24</sup>The range of the risk-aversion parameter varies wildly in studies that try to match asset-pricing facts. For example, [Tallarini \(2000\)](#) uses a risk aversion parameter above 50 in order to generate a high market price of risk, whereas [Bansal and Yaron \(2004\)](#) use low values of risk aversion in environments with long-run risks and stochastic volatility. Note that the plausibility of the size of atemporal risk aversion cannot be judged independently from the stochastic processes that drive uncertainty in the economy, since they jointly bear implications for the premium for early resolution of uncertainty. See [Epstein et al. \(2014\)](#) for a thoughtful evaluation of calibration practices in the asset-pricing literature from this angle.

<sup>25</sup>Note that if the initial shock was high,  $g_0 = g_H$ , we would have  $b_H = 0$  and  $b_L > 0$ . The planner insures against adverse shocks by running a deficit when government expenditures are high, which is financed by debt contingent on a low expenditure shock. When shocks are low, the planner runs a surplus to pay back the issued debt.

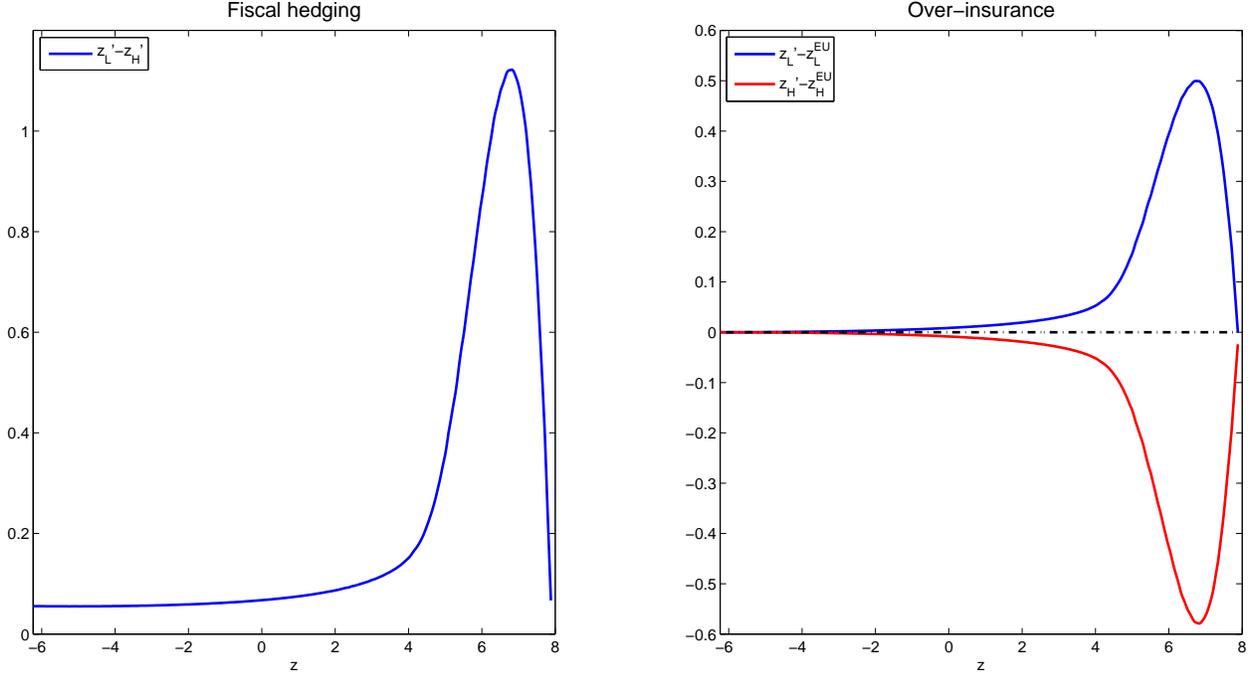


Figure 1: The left panel depicts the difference  $z'_L - z'_H$ . The difference starts decreasing at high values of  $z$ , because the probability of a binding upper bound increases. The right panel compares positions for recursive and time-additive utility. For both graphs the current shock is low,  $g = g_L$ . A similar picture emerges when  $g = g_H$ .

government hedges fiscal shocks by issuing more debt in marginal utility units for the low shock and less for the high shock,  $z'_L > z'_H$ . Thus, as highlighted in the overview of the mechanism, propositions 2 and 5 imply that tax distortions decrease when fiscal shocks are high,  $\Phi'_L > \Phi > \Phi'_H$  and  $\tau'_L > \tau > \tau'_H$ . The right panel in figure 1 plots the difference in the policy functions in the recursive utility and the expected utility case,  $z'_i - z_i^{\text{EU}}, i = L, H$ , in order to demonstrate the “over-insurance” property of the optimal plan: against  $g_L$ , the planner is issuing more debt than he would in the time-additive economy. Similarly, debt against  $g_H$  is less than its respective value in an economy where  $\rho = \gamma$ . So the planner is actively taking larger positions in absolute value.<sup>26</sup>

To see the price manipulation that takes place with recursive utility, figure 2 contrasts the optimal stochastic discount factor  $S(g' = g_i, z, g), i = L, H$ , (top and bottom left panels), to the induced stochastic discount factor that pertains to a sub-optimal constant- $\Phi$  policy, that ignores the beneficial pricing effects of continuation values (top and bottom right panels). By contrasting the left to the right panels, we see how the planner, by issuing more debt against  $g_L$  and increasing the respective tax rate, manages to increase the pricing kernel and therefore the price of a claim to consumption, making debt cheaper. Note that the increase in the stochastic discount factor due to the continuation value part is naturally reinforced by an increase in the period marginal utility

<sup>26</sup>A virtually identical graph would emerge if we compared the optimal policy functions  $z'_i$  with the positions that would be induced in a recursive utility economy with a planner that follows a sub-optimal, constant excess burden policy.

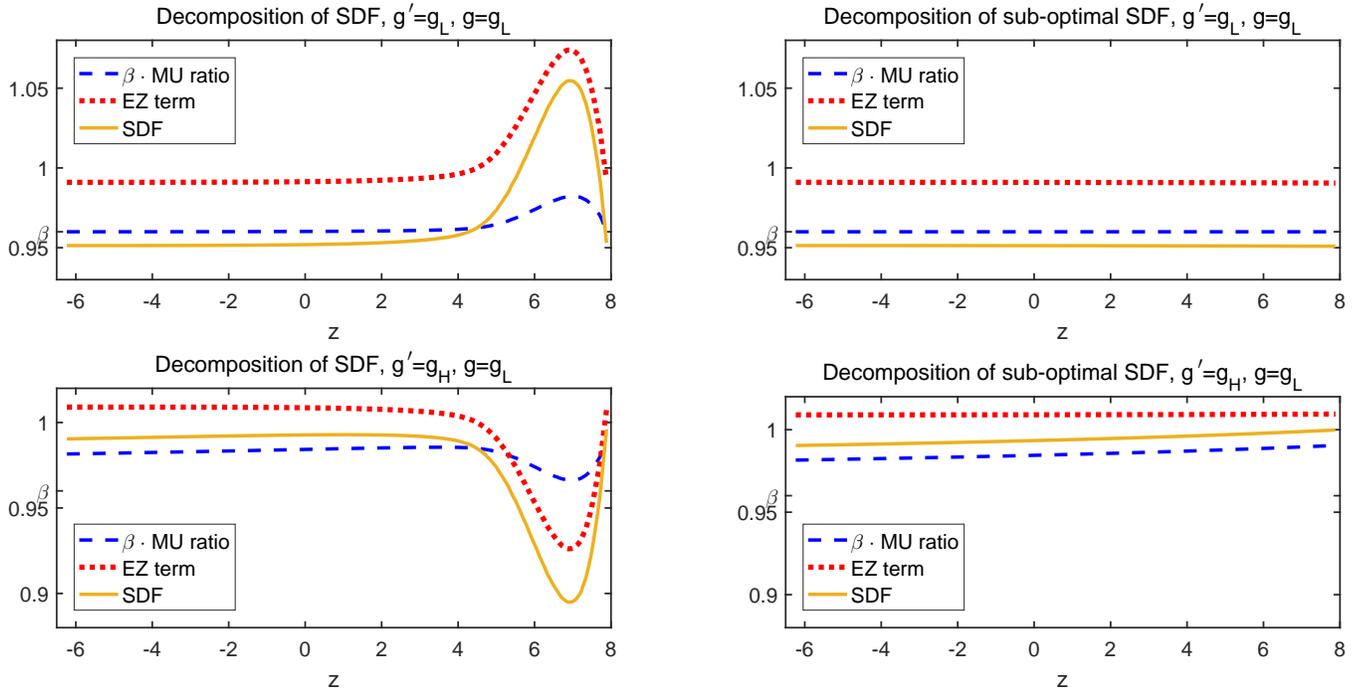


Figure 2: The left panels decompose the optimal stochastic discount factor to its period marginal utility and continuation value part, when the current policy shock is low,  $g = g_L$ . The right panels perform the same exercise assuming that a sub-optimal, constant- $\Phi$  policy is followed. A similar picture emerges when  $g = g_H$ .

part due to decreased future consumption. Similarly, by issuing less debt or buying more assets against a high fiscal shock, and taxing consequently less, the planner is decreasing the pricing kernel for bad states of the world.

## 6.5 Persistence and negative correlation with spending

Consider a simulation of 10,000 sample paths that are 2,000 periods long. Table 1 highlights the persistence that propositions 3 and 5 hinted at. The median persistence of the tax rate is very high (0.998), despite the fact that government expenditure shocks are i.i.d., which contrasts to the standard history-independence result of Lucas and Stokey (1983). As expected, the *change* in the tax rates is strongly negatively correlated with government expenditures ( $-0.99$ ) and therefore with output.<sup>27</sup>

## 6.6 Back-loading and volatility of distortions

Figure 3 plots the mean, standard deviation, the 5th and the 95th percentile of the tax rate and the debt-to-output ratio. It shows that there is a *positive* drift in the tax rate with respect to the

<sup>27</sup>The theory predicts that changes in tax rates are affected by the relative debt position, which is highly negatively correlated with fiscal shocks. In contrast, the level of the tax rate is affected by the cumulative relative debt position  $\sum_i \eta_i$ , leading overall to a small correlation with government expenditures.

Table 1: Statistics of tax rate sample paths.

	<i>Recursive utility</i>	
	short samples	long samples
Autocorrelation	0.9791	0.9980
Correlation of $\Delta\tau$ with $g$	-0.9999	-0.9984
Correlation of $\Delta\tau$ with output	-0.9977	-0.9762
Correlation of $\tau$ with $g$	-0.1098	-0.0346
Correlation of $\tau$ with output	-0.1793	-0.2418

The table reports median sample statistics across 10,000 sample paths of the tax rate. For the time-additive case the respective moments are not well defined since the tax rate is constant. For the recursive utility case the median sample statistics are calculated for short samples (the first 200 periods) and long samples (2,000 periods).

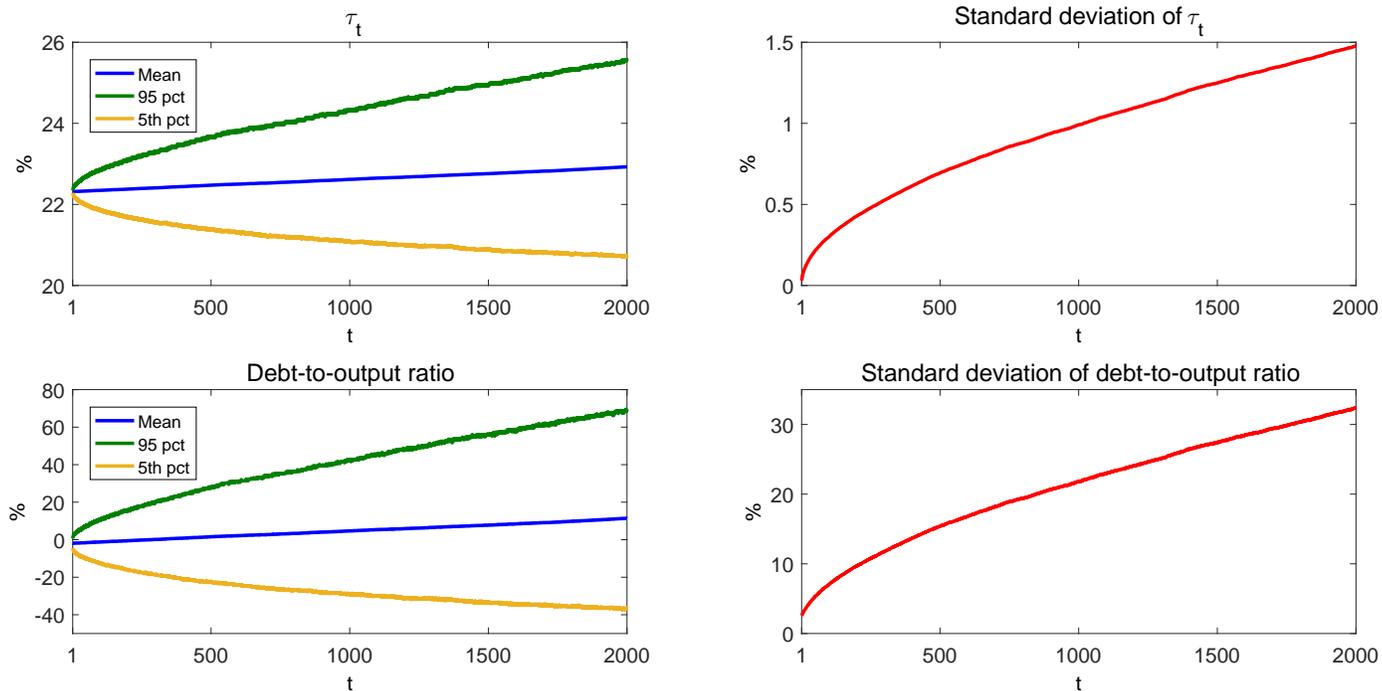


Figure 3: Ensemble moments of the tax rate and the debt-to-output ratio.

data-generating process, which is mirrored also in the debt-to-output ratio. This back-loading of distortions reflects the submartingale results of propositions 3 and 5. The increase in the mean tax rate is slow (about 60 basis points in 2,000 periods) but the standard deviation rises to almost 1.5 percentage points. So the distribution of the tax rate is “fanning-out” over time. Similarly, the mean and the standard deviation of the debt-to-output rise to 11 and 32 percentage points

Table 2: Moments from the stationary distribution.

<b>Stationary distribution</b>					
		$\tau$ in %	$b/y$ in %		
Mean		30.86	181.97		
St. dev.		4.94	104.28		
98th pct		40.6	397.3		
St. dev. of change		0.17	12.72		
Autocorrelation		0.9994	0.9926		
<b>Correlations</b> ( $\tau, b, g$ )					
Corr( $\Delta\tau, g$ )	-0.6183	Corr( $\Delta b, g$ )	-0.7639	Corr( $\Delta\tau, b$ )	0.0476
Corr( $\Delta\tau, \Delta g$ )	-0.4383	Corr( $\Delta b, \Delta g$ )	-0.9070	Corr( $\Delta\tau, \Delta b$ )	0.7228
Corr( $\tau, g$ )	-0.0219	Corr( $b, g$ )	-0.0653	Corr( $\tau, b$ )	0.9933

The simulation is 60 million periods long. The first 2 million periods were dropped. Remember that in the expected utility case the tax rate is 22.3% and that the debt-to-output ratio has mean -1.91% and a standard deviation of 1.91%.

respectively at  $t = 2,000$ .<sup>28</sup>

## 6.7 Long-run

The martingale property of the inverse of the excess burden may introduce non-stationarities in the long-run. The asymptotic behavior of taxes and debt depends on two objects: the behavior of the relative debt position  $\eta_{t+1}$  in the long-run and the upper bounds of the state space. For example, if the relative debt position converges to zero, then the excess burden, and therefore the tax rate, would converge to a constant. Furthermore, if there is always back-loading of taxes with respect to the physical measure (which is not necessarily the case since propositions 3 and 5 involve  $\pi_t \cdot M_t$ ), there will be progressively high accumulation of debt and at some point fiscal hedging may become limited, due to an upper bound on debt issuance.

Recall that the proper state variable of the commitment problem is debt in marginal utility units. Consequently, even if there is a natural upper bound in terms of debt (the maximal present discounted value of surpluses), there may not be an upper bound in terms of debt in marginal utility units. To see that, consider a situation where the tax rate is so large that consumption decreases to zero. Then marginal utility goes to infinity and debt in marginal utility units *may* inherit the same behavior.<sup>29</sup> Computation obviously requires an upper bound. If this is occasionally binding,

<sup>28</sup>See the Online Appendix for additional simulations with either higher risk aversion or higher shock volatility.

<sup>29</sup>The exact behavior depends heavily on the upper bounds of the surplus in marginal utility units,  $U_c c - U_l h$ . See the Online Appendix.

the positive drift of the tax rate breaks down and its distribution becomes stationary.<sup>30</sup>

For the particular period utility function of the quantitative exercise, I prove in the Online Appendix that there are no positive convergence points for  $\Phi_t$  (which concern essentially the asymptotic behavior of  $\eta_{t+1}$ ). Since this is the case, my choices on the size of the state space are driven by computational considerations. The computational exercise has upper bounds that correspond to a debt-to-output ratio close to 600%.<sup>31</sup> Table 2 reports moments of interest from the stationary distribution. The tax rate has mean 30.8% and standard deviation close to 5 percentage points. This tax rate is pretty high: it supports debt-to-output ratios that have mean 182% with a standard deviation of 105 percentage points. The conditional volatility of the tax rate and the debt-to-output ratio are small but the unconditional volatility is large due to the extremely high persistence in the long-run.<sup>32</sup>

## 7 Optimal debt returns and fiscal insurance

In this section I am taking a deeper look at the theory of debt management with recursive utility. I focus on the use of the return of the government debt portfolio as a tool of fiscal insurance.<sup>33</sup> To that end, I measure *optimal* fiscal insurance in simulated data by using the decomposition of [Berndt et al. \(2012\)](#) (BLY henceforth) and contrast it to their empirical findings.

BLY devised a method to quantify fiscal insurance in post-war US data by log-linearizing the intertemporal budget constraint of the government.<sup>34</sup> Let the government budget constraint be

---

<sup>30</sup>What breaks down is the martingale result of proposition 3. The optimality condition with respect to  $z$  when there is an upper bound on  $z$  becomes  $\Phi_{t+1}(1 + (1 - \beta)(1 - \gamma)\eta_{t+1}\Phi_t) \leq \Phi_t$ . If  $1 + (1 - \beta)(1 - \gamma)\eta_{t+1}\Phi_t > 0$ , we get  $\frac{1}{\Phi_{t+1}} \geq \frac{1}{\Phi_t} + (1 - \beta)(1 - \gamma)\eta_{t+1}$ , which implies that  $1/\Phi_t$  is a submartingale (and *not* a martingale) with respect to  $\pi_t \cdot M_t$ . Therefore, the convexity of function  $f$  in the proof of proposition 3 is not sufficient anymore to infer  $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$  (we need also  $f$  to be monotonically increasing and it is actually decreasing). The same reasoning applies to the tax rate in proposition 5.

<sup>31</sup>The larger the upper bounds, the larger the non-convexities associated with recursive utility, which lead to non-convergence issues. It turns out that the particular upper bounds are rarely visited (the 98th percentile of the debt-to-output ratio is about 400%). This is a numerical statement that may not hold for other parameterizations. In the Online Appendix I provide several robustness exercises with respect to the size of the state space. Furthermore, I consider different period utility functions, for which the existence of a stationary distribution is more probable, without having to rely on ad-hoc upper bounds.

<sup>32</sup>In the Online Appendix I provide instructive sample paths and moments from a persistent shock specification: I use the government spending shocks of [Chari et al. \(1994\)](#) (see also next section). There are two big differences with persistent shocks: first, the unconditional volatility is similar to the baseline case, but the volatility of the change of the tax rate ( $\Delta\tau$ ) is more than doubled. Second, the speed at which the stationary distribution is reached is much higher. The mean and the standard deviation of the tax rate increase by 4 and 5 percentage points respectively in 2,000 periods (in contrast to the lower medium-run numbers displayed in figure 3). Overall, there is more action both in the tax rate and in debt when shocks are persistent.

<sup>33</sup>See [Hall and Sargent \(2011\)](#) for the careful measurement of the return of the government debt portfolio and [Hall and Krieger \(2000\)](#) for an analysis of optimal debt returns in the [Lucas and Stokey \(1983\)](#) setup. [Marcet and Scott \(2009\)](#) contrast fiscal insurance in complete and incomplete markets.

<sup>34</sup>Their exercise follows the spirit of [Campbell \(1993\)](#) – who worked with the household’s budget constraint – and [Gourinchas and Rey \(2007\)](#) – who employed the country’s external constraint.

Table 3: Returns on government debt portfolio,  $R(g', g, z)$ .

$\mathbf{R} - \mathbf{1}$ in %	Expected utility at $z_{\text{EU}}$		Recursive utility at $z_{\text{EU}}$		Recursive utility at $E(z)$	
	$g_L$	$g_H$	$g_L$	$g_H$	$g_L$	$g_H$
$g_L$	5.96	-27.92	6.95	-41.15	5.30	-15.61
$g_H$	49.47	1.68	69.52	0.81	25.96	3.10

Rows denote current shock. The value of  $z$  in the expected utility case is  $(z_L, z_H) = (0.7795, 0.5399)$ . The average value of  $z$  with recursive utility is  $E(z) = 2.2626$ .

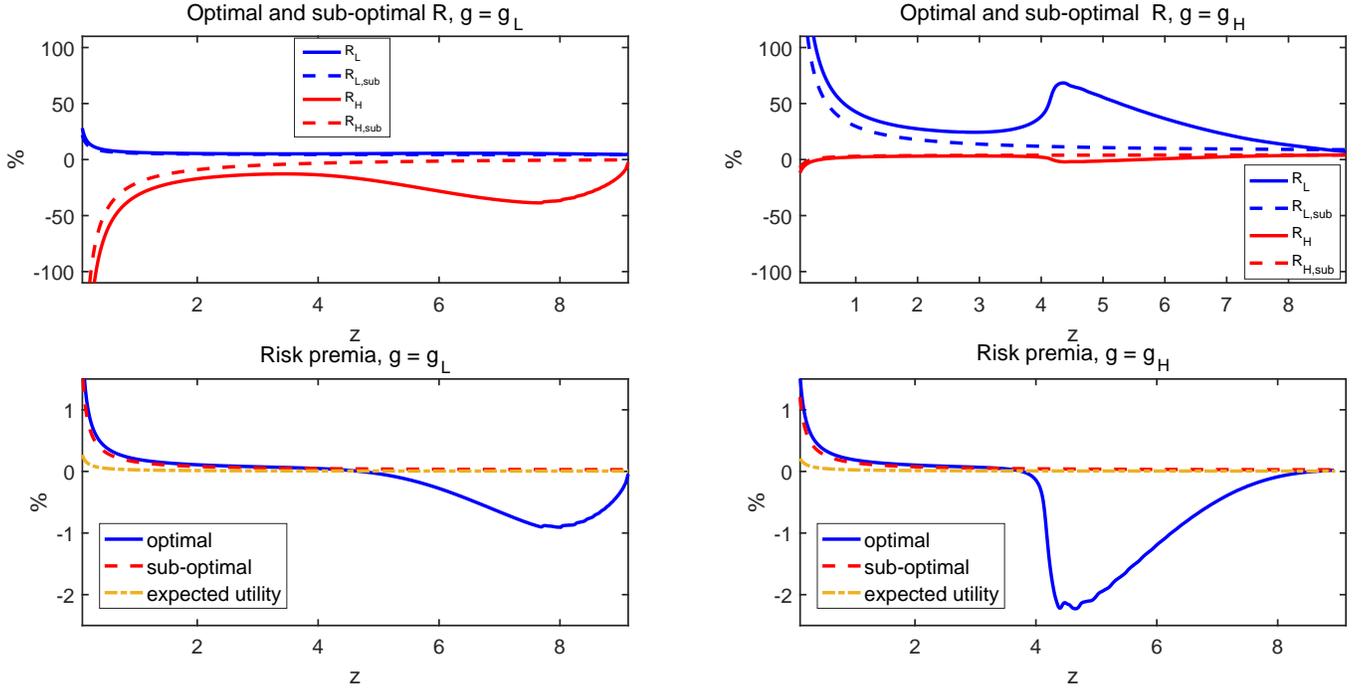


Figure 4: The top panels contrast the optimal  $R(g', g, z)$  to the sub-optimal return coming from a constant  $\Phi$  policy. The bottom panels plot the respective conditional risk premia, where I also include the expected utility risk-premia for comparison.

written as

$$b_{t+1} = R_{t+1} \cdot (b_t + g_t - T_t), \quad (27)$$

where  $R_{t+1} \equiv b_{t+1}(g^{t+1}) / \sum_{g_{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1})$ , the return on the government debt portfolio, constructed in the model economy by the state-contingent positions  $b_{t+1}$ , and  $T_t \equiv \tau_t h_t$ , the tax revenues. By construction, we have  $\sum_{g_{t+1}} p_t(g_{t+1}, g^t) R_{t+1}(g^{t+1}) = 1$ . BLY log-linearize (27) and derive a representation in terms of *news* or *surprises* in the present value of government

expenditures, returns and tax revenues,<sup>35</sup>

$$I_{t+1}^g = -\frac{1}{\mu_g} I_{t+1}^R + \frac{1}{\mu_g} I_{t+1}^T, \quad (28)$$

where

$$\begin{aligned} I_{t+1}^g &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{BLY}}^i \Delta \ln g_{t+i+1} \\ I_{t+1}^R &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{BLY}}^i \ln R_{t+i+1} \\ I_{t+1}^T &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{BLY}}^i \mu_T \Delta \ln T_{t+i+1}, \end{aligned} \quad (29)$$

and  $(\mu_g, \mu_T, \rho_{\text{BLY}})$  approximation constants. Decomposition (28) captures how a fiscal shock is absorbed: a positive surprise in the growth rate of spending,  $I_{t+1}^g$ , is financed by either a negative surprise in (current or future) returns,  $I_{t+1}^R$ , or by a positive surprise in (current or future) growth rates of tax revenues,  $I_{t+1}^T$ . BLY refer to these types of fiscal adjustment as the *debt valuation* channel and the *surplus* channel respectively. The decomposition can be written in terms of fiscal adjustment *betas*,

$$1 = -\frac{\beta_R}{\mu_g} + \frac{\beta_T}{\mu_g}, \quad \text{where} \quad \beta_R \equiv \frac{\text{Cov}(I_{t+1}^g, I_{t+1}^R)}{\text{Var}(I_{t+1}^g)}, \quad \beta_T \equiv \frac{\text{Cov}(I_{t+1}^g, I_{t+1}^T)}{\text{Var}(I_{t+1}^g)}. \quad (30)$$

The fraction of fiscal shocks absorbed by debt returns and tax revenues are  $-\beta_R/\mu_g$  and  $\beta_T/\mu_g$  respectively. *Fiscal insurance* refers to the reduction of returns in light of a positive fiscal shock,  $\beta_R < 0$ .

## 7.1 Returns and risk premia

For the fiscal insurance exercise I use the Chari et al. (1994) specification of fiscal shocks that captures well the dynamics of government consumption in post-war U.S. data. I set initial debt to 50% of first-best output. The utility function and the calibration of the rest of the parameters is the same as in the previous section.

Table 3 provides the conditional returns of the government debt portfolio implied by the Ramsey plan in the expected and recursive utility economy. It shows the essence of debt return manage-

---

<sup>35</sup>See Berndt et al. (2012) for the derivations and the Online Appendix for the definition of the approximation constants.

ment, i.e. the reduction of the return on government debt in bad times in exchange of an increase in return in good times. For example, in the expected utility economy bond-holders suffer capital losses of -28% when there is a switch from a low to a high fiscal shock. They still buy government debt because they are compensated with a high return of 49% when there is a switch back to a low shock. The gains and losses to the bond-holders at the same value of the state variable  $z$  with recursive utility are much larger (-41% and 69% respectively), due to the “over-insurance” property. On average though, the government issues larger quantities of debt with recursive utility, which actually makes the size of conditional returns necessary to absorb fiscal shocks *smaller*. This can be seen in the third part of table 3, which displays the conditional returns for EZW utility at the average debt holdings,  $E(z)$ .

Figure 4 takes a closer look at the returns of the government portfolio. The top panels demonstrate the desire of the government to increase the returns of the debt portfolio for good shocks and decrease it for bad shocks, by contrasting the optimal returns with recursive utility with the sub-optimal returns that are induced by a constant- $\Phi$  policy. The bottom panels plot the conditional premium of government debt over the risk-free rate for recursive utility (following either optimal or sub-optimal policy) and for expected utility. What is interesting to observe is the fact that for large levels of debt, when over-insurance becomes even more pronounced, the optimal conditional risk premium of government debt becomes *negative*.<sup>36</sup>

The reason for government debt becoming a *hedge* is simple: the risk premium over the risk-free rate  $R_t^F$  can be expressed as  $E_t R_{t+1}/R_t^F - 1 = -Cov_t(S_{t+1}, R_{t+1})$ . Debt returns are high when fiscal shocks are low. But optimal policy with recursive utility prescribes large tax rates at exactly these states of the world. As a result, at some point tax rates at good shocks become so high that both consumption and continuation values of agents *fall* (despite shocks being favorable), and therefore  $S_{t+1}$  increases. This leads to a positive covariance of the stochastic discount factor with government returns and a negative risk premium. In other words, optimal policy converts “good” times (with low  $g$ ) to “bad” times with high tax rates (and “bad” times with high  $g$  to “good” times with low tax rates). Thus, the household is happy to accept a negative risk premium for a security that pays well when tax rates are so high.<sup>37</sup>

## 7.2 Fiscal insurance

Table 4 reports the correlations and the standard deviations of news to government spending, debt returns and tax revenues at the stationary distribution and table 5 reports the respective fiscal adjustment betas and fiscal insurance fractions. For both the expected and the recursive utility

<sup>36</sup>The unconditional risk premium remains positive. See the Online Appendix for additional information on average returns and the market price of risk.

<sup>37</sup>We can see also the change in the ranking of the discount factors in figure 2. For large enough debt we have  $S(g_L, z, g) > S(g_H, z, g)$ , whereas for low enough debt the opposite holds. With either expected utility, or a sub-optimal constant  $\Phi$  policy and recursive utility, this does not happen and we always have  $S(g_L, z, g) < S(g_H, z, g)$  and a positive conditional risk premium.

Table 4: News to expenditures, returns and revenues.

	Expected utility			Recursive utility		
	$I^g$	$I^R$	$I^T$	$I^g$	$I^R$	$I^T$
$I^g$	0.91			0.98		
$I^R$	-1	8.60		-0.79	7.52	
$I^T$	1	-1	1.37	-0.74	0.53	2.44

Standard deviations (on the diagonal, multiplied by 100) and correlations of the news variables at the stationary distribution. Calibration of shocks as in [Chari et al. \(1994\)](#).

Table 5: Fiscal insurance.

	Expected utility		Recursive utility	
	Valuation channel	Surplus channel	Valuation channel	Surplus channel
<b>Beta</b>	-9.46	1.51	-6.05	-1.85
Current	-9.65	5.28	-6.20	-0.64
Future	0.19	-3.77	0.15	-1.21
<b>Fraction in %</b>	87.85	13.99	180.93	-55.16
Current	89.67	49.08	185.13	-19.13
Future	-1.82	-35.09	-4.20	-36.03

Fiscal adjustment betas and fiscal insurance fractions. The approximation constants are  $(\mu_g, \mu_T, \rho_{BLY}) = (10.7654, 11.7654, 0.958)$  and  $(\mu_g, \mu_T, \rho_{BLY}) = (3.3462, 4.3462, 0.9525)$  in the time-additive and recursive utility case respectively. The  $R^2$  in the expected utility case is almost 100% for both regressions. For the recursive utility economy the  $R^2$  is 62.44% and 55.07% for the return and revenues regression respectively. The current return beta comes from regressing  $\ln R_{t+1} - E_t \ln R_{t+1}$  on news to spending. Similarly, the current tax revenue beta comes from regressing current news to the growth in tax revenues on news to spending.

case news to optimal returns are pretty volatile and negatively correlated to fiscal shocks. What is important to notice is that news to the growth rate in tax revenues ( $I^T$ ) are positively correlated with news to fiscal shocks in the expected utility case (absorbing therefore part of the fiscal shock) but *negatively* correlated in the recursive utility case.

Turning to fiscal insurance fractions, about 87% of fiscal risk is absorbed by the debt valuation channel and about 13% by the surplus channel in the expected utility economy. Thus, the debt valuation channel is prominent in the absorption of shocks. Fiscal insurance motives are *amplified* with recursive utility: the planner is reducing even more returns in the face of adverse shocks, to the point where the tax rate is actually *reduced*, explaining the negative correlation we saw in [table 4](#). As a result, the reliance on the debt valuation channel is even larger and the surplus channel becomes essentially *inoperative*. The fraction of fiscal risk absorbed by reductions in the market value of debt is about 180% (predominantly by a reduction in *current* returns), which allows the

government to reduce the growth in tax revenues, leading to a surplus channel of  $-55\%$ .<sup>38</sup>

**Is actual fiscal insurance even worse than we thought?** BLY measure fiscal insurance on post-war U.S. data. They focus on defense spending in order to capture the exogeneity of government expenditures and show that 9% of defense spending shocks has been absorbed by a reduction in returns (mainly through future returns) and 73% by an increase in non-defense surpluses. Thus, there is some amount of fiscal insurance in the data; smaller though than what optimal policy in an expected utility economy would recommend. The current exercise shows that in environments that can generate a higher market price of risk, governments debt returns have to be used to a much *greater* extent as a fiscal shock absorber. Thus, if we were to evaluate actual fiscal policy through the normative prescriptions of the recursive utility economy, the following conclusion emerges: actual fiscal policy is even *worse* than we thought.

## 8 Economy with capital

Consider now an economy with capital as in [Zhu \(1992\)](#) and [Chari et al. \(1994\)](#) and recursive preferences. Let  $s$  capture uncertainty about government expenditure or technology shocks, with the probability of a partial history denoted by  $\pi_t(s^t)$ . The resource constraint in an economy with capital reads

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)), \quad (31)$$

where  $\delta$  denotes the depreciation rate,  $k_{t+1}(s^t)$  capital measurable with respect to  $s^t$  and  $F$  a constant returns to scale production function. The representative household accumulates capital, that can be rented at rental rate  $r_t(s^t)$ , and pays capital income taxes with rate  $\tau_t^K(s^t)$ . The household's budget constraint reads

$$c_t(s^t) + k_{t+1}(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}) \leq (1 - \tau_t(s^t)) w_t(s^t) h_t(s^t) + R_t^K(s^t) k_t(s^{t-1}) + b_t(s^t),$$

---

<sup>38</sup>The fractions do not add to 100% due to the approximation error coming from log-linearizing (27). The same issue emerges with actual fractions from post-war U.S. data (see [Berndt et al. \(2012\)](#) and the respective table in the Online Appendix that reproduces their results). Furthermore, two robustness exercises are provided in the Online Appendix. At first, in order to apply the log-linear methodology of [Berndt et al. \(2012\)](#), I excluded negative debt realizations that amount to 4.4% of the stationary distribution. In the Online Appendix I use a linear approximation of (27) that allows me to include this type of observations. The size of the valuation and surplus channel for both expected and recursive utility remains essentially the same. Second, one may think that the stark contrast between expected and recursive utility is coming from the much larger debt and taxes in the latter case, a fact which is reflected in the very different approximation constants across the two economies. In order to control that, I calculate in the Online Appendix the expected utility fiscal insurance fractions by setting initial debt equal to the mean of the recursive utility economy. This leads to similar approximation constants with the recursive utility case, so any difference in the fiscal channels is stemming from the endogenous reaction of returns and tax revenues. The valuation and surplus channel in the expected utility economy become 83% and 17% respectively, so the difference between expected and recursive utility is even starker.

where  $R_t^K(s^t) \equiv (1 - \tau_t^K(s^t))r_t(s^t) + 1 - \delta$ , the after-tax gross return on capital.

I provide the details of the competitive equilibrium and the analysis of the Ramsey problem in the Appendix and summarize here the main results. In short, the completeness of the markets allows the recasting of the household's budget constraint in terms of *wealth*,  $W_t \equiv b_t + R_t^K k_t$ , making therefore wealth in marginal utility units,  $z_t \equiv U_{ct}W_t$ , the relevant state variable for the optimal taxation problem. With this interpretation of  $z_t$ , the dynamic implementability constraint remains the same as in an economy without capital. The recursive formulation of the Ramsey problem has  $(z, k, s)$  as state variables. The excess burden of taxation  $\Phi$  captures now the shadow cost of an additional unit of wealth in marginal utility units,  $\Phi = -v_z(z, k, s)$ , where  $v$  denotes the value function. As expected, the excess burden of taxation is not constant anymore across states and dates. In particular, we have:

**Proposition 6.** *The law of motion of  $\Phi_t$  in an economy with capital remains the same as in (19), with  $\eta_{t+1}$  defined as in (18), denoting now the relative wealth position in marginal utility units, with an average of zero,  $E_t m_{t+1} \eta_{t+1} = 0$ . Let  $\hat{s}$  and  $\tilde{s}$  denote states of the world at  $t + 1$  for which  $\eta_{t+1}(\hat{s}) > 0 > \eta_{t+1}(\tilde{s})$ . Then  $\Phi_{t+1}(\hat{s}) > \Phi_t > \Phi_{t+1}(\tilde{s})$ , when  $\rho < \gamma$ . Propositions 3, 4 and 5 go through, so the same conclusions are drawn for the dynamics of the excess burden and the labor tax as in an economy without capital.*

Proposition 6 generalizes our previous results about the excess burden of taxation and the labor tax. Recall that in an economy without capital the planner was taxing more events against which he was issuing relatively more debt in order to take advantage of the positive covariance between debt in marginal utility units and the stochastic discount factor, through the channel of continuation values. Market completeness makes state-contingent wealth in marginal utility units the relevant hedging instrument in an economy with capital. Note also that we allowed technology shocks in the specification of uncertainty in this section, in addition to the typical government expenditure shocks. We expect that the planner hedges adverse shocks, which are high fiscal shocks and low technology shocks with low wealth positions, and favorable shocks, i.e. low fiscal shocks or high technology shocks with high wealth positions. If this is the case, the planner decreases the labor tax for high spending shocks and low technology shocks, mitigating again the effects of shocks. The opposite happens for favorable shocks.

## 8.1 Capital taxation

Capital accumulation affects through continuation values the pricing of state-contingent claims, a fact which alters the incentives for taxation at the intertemporal margin. In particular, the optimal accumulation of capital is governed by (details in the Appendix),

$$E_t S_{t+1}^* (1 - \delta + F_{K,t+1}) = 1, \quad \text{where } S_{t+1}^* \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{\lambda_{t+1}/\Phi_{t+1}}{\lambda_t/\Phi_t}, \quad (32)$$

where  $\lambda_t$  stands for the multiplier on the resource constraint (31) in the recursive formulation of the second-best problem.

I call  $S_{t+1}^*$  the *planner's* stochastic discount factor. The discount factor  $S_{t+1}^*$  captures how the planner discounts the *pre-tax* capital return  $1 - \delta + F_{K,t+1}$  at the *second-best* allocation.  $S_{t+1}^*$  contrasts to the market stochastic discount factor,  $S_{t+1} \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}/U_{c,t}$ , which prices *after-tax* returns,  $E_t S_{t+1} R_{t+1}^K = 1$ . In a first-best world with lump-sum taxes available, we identically have  $S_{t+1}^* \equiv S_{t+1}$ . At the second-best, the difference in the two discount factors  $S_{t+1} - S_{t+1}^*$  is useful in summarizing the optimal wedge at the intertemporal margin, in the form of the *ex-ante* tax rate on capital income.

In particular, as is well known from Zhu (1992) and Chari et al. (1994), only the non-state contingent ex-ante capital tax  $\bar{\tau}_{t+1}^K(s^t)$  can be uniquely determined by the second-best allocation. This tax is defined as  $\bar{\tau}_{t+1}^K \equiv (E_t S_{t+1} (1 - \delta + F_{K,t+1}) - 1) / E_t S_{t+1} F_{K,t+1}$ , which by (32) becomes

$$\bar{\tau}_{t+1}^K = \frac{E_t [S_{t+1} - S_{t+1}^*] (1 - \delta + F_{K,t+1})}{E_t S_{t+1} F_{K,t+1}}. \quad (33)$$

Thus, the sign of the ex-ante capital tax is determined by the numerator in (33), i.e. the non-centered covariance of the two discount factors with the pre-tax capital return. The difference  $S_{t+1} - S_{t+1}^*$  can be expressed in terms of differences in the inverse of the excess burden of taxation and differences in the own and cross elasticity of the marginal utility of consumption, which leads to the following proposition about capital taxation.<sup>39</sup>

**Proposition 7.** (*Capital taxation criterion*) *The ex-ante tax rate on capital income  $\bar{\tau}_{t+1}^K, t \geq 1$  is positive (negative) iff*

$$E_t \zeta_{t+1} \left[ \underbrace{\left( \frac{1}{\Phi_t} - \frac{1}{\Phi_{t+1}} \right)}_{\text{change in } 1/\Phi_t} + \underbrace{\left( \epsilon_{cc,t+1} + \epsilon_{ch,t+1} - \epsilon_{cc,t} - \epsilon_{ch,t} \right)}_{\text{change in period elasticities}} \right] > (<) 0,$$

with weights  $\zeta_{t+1} \equiv S_{t+1} (1 - \delta + F_{K,t+1}) / E_t S_{t+1} (1 - \delta + F_{K,t+1})$ . If  $\epsilon_{cc} + \epsilon_{ch}$  is constant, then any capital taxation comes from variation in the excess burden  $\Phi_t$ .

*Proof.* See Appendix. □

The ex-ante capital tax furnishes by construction the same present discounted value of tax

<sup>39</sup>As it was the case with the labor tax in footnote 19, the capital tax criterion applies also for the deterministic and stochastic time-additive case for any standard  $U$ .

revenues as any vector of feasible state-contingent capital taxes. As such, it *averages* intertemporal distortions across states next period, with weights  $\zeta_{t+1}$  that depend on the stochastic discount factor and the pre-tax capital return. The distortions at each state next period depend on *both* the change in the elasticity of the marginal utility of consumption (the time-additive part) *and* the change in the excess burden of taxation (the novel recursive utility part).

**Time-additive economy.** Assume that we are either in a deterministic economy or in a stochastic but time-additive economy with  $\rho = \gamma$ . In both cases  $\Phi_t$  is constant and the capital taxation criterion of proposition 7 depends only on the change in period elasticities. For the deterministic case, capital income is taxed (subsidized) if the sum of the own and cross elasticity is increasing (decreasing). A necessary and sufficient condition for a zero capital tax at *every* period from period two onward is a *constant* sum of elasticities,  $\epsilon_{cc} + \epsilon_{ch}$ , which implies that  $S_{t+1}^* = S_{t+1}$ . If the period utility function is such so that the elasticities are *not* constant for each period, then there is zero tax on capital income only at the deterministic *steady state*, where the constancy of the consumption-labor allocation delivers constant elasticities. This delivers the steady-state results of Chamley (1986) and Judd (1985). In the stochastic case of Chari et al. (1994) and Zhu (1992), the sign of the ex-ante capital tax depends on the *weighted* average of the change in elasticities.<sup>40</sup>

**Recursive utility.** The full version of the capital tax criterion in proposition 7 applies when  $\rho \neq \gamma$ . To focus on the novel effects of recursive utility, consider the case of constant period elasticities and assume that  $\rho < \gamma$ . For an example in this class, let the composite good be

$$u(c, 1 - h) = [c^{1-\rho} - (1 - \rho)v(h)]^{\frac{1}{1-\rho}}, \quad v', v'' > 0, \quad (34)$$

that delivers a period utility  $U = (u^{1-\rho} - 1)/(1 - \rho)$ , which is separable between consumption and leisure and isoelastic in consumption.<sup>41</sup> Chari et al. (1994) and Zhu (1992) have demonstrated that these preferences deliver a *zero* ex-ante capital tax from period two onward. This is easily interpreted in terms of proposition 7, since  $\epsilon_{cc} = \rho$  and  $\epsilon_{ch} = 0$ .

With recursive preferences though, even in the constant period elasticity case, there is a novel source of taxation coming from the willingness of the planner to take advantage of the pricing effects of continuation values. By using the law of motion of the excess burden of taxation (19) to substitute  $\eta_{t+1}$  for the change in  $1/\Phi_t$ , the criterion becomes

$$\bar{\tau}_{t+1}^K > (<) 0 \quad \text{iff} \quad E_t \zeta_{t+1} \eta_{t+1} > (<) 0, \quad \text{when } \rho < \gamma. \quad (35)$$

---

<sup>40</sup>Variation in  $\epsilon_{cc} + \epsilon_{ch}$  is a necessary condition for a non-zero ex-ante capital tax, but is not sufficient anymore since the weighted average can still in principle deliver a zero tax.

<sup>41</sup>The same comments as in footnote 22 apply. The constant Frisch elasticity case is obviously a member of this class.

Thus, the capital taxation criterion depends on the weighted average of the relative wealth positions  $\eta_{t+1}$ . To understand the logic behind the criterion, note that the *change* in the excess burden of taxation determines the sign of distortions at the intertemporal margin. States where there are positive relative wealth positions ( $\eta_{t+1} > 0$ ), make the planner increase the excess burden of taxation,  $\Phi_{t+1} > \Phi_t$ . This raises the labor tax and leads to a planner's discount factor that is smaller than the market discount factor,  $S_{t+1}^* < S_{t+1}$ , which we can think of as introducing a *state-contingent* capital tax.<sup>42</sup> To understand the intuition, a positive state-contingent capital tax reduces capital accumulation and therefore utility. In a recursive utility world this increases the price of the respective Arrow claim and the value of state-contingent wealth. This appreciation of the value of wealth is beneficial when wealth positions are relatively large ( $\eta_{t+1} > 0$ ). In the opposite case of  $\eta_{t+1} < 0$  the planner is decreasing the labor tax and has the incentive to put a state-contingent capital subsidy ( $S_{t+1}^* > S_{t+1}$ ). The ex-ante capital tax depends on the weighing of the positive versus the negative intertemporal distortions.

## 8.2 Ex-ante subsidy

To gain more insight about the sign of the ex-ante capital tax, we need to understand the behavior of the weights  $\zeta_{t+1}$ . Consider the separable preferences in (34) and let  $\rho = 1 < \gamma$ . Then, by using the property that  $E_t m_{t+1} \eta_{t+1} = 0$  and the definition of  $\zeta_{t+1}$ , the capital tax criterion simplifies to

$$\bar{\tau}_{t+1}^K > (<) 0 \quad \text{iff} \quad \text{Cov}_t^M(c_{t+1}^{-1} \cdot (1 - \delta + F_{K,t+1}), z_{t+1}) > (<) 0.$$

Thus, we can express the criterion in terms of the conditional *covariance* (with respect to the continuation-value adjusted measure M) of the marginal utility weighted pre-tax capital return with the wealth positions in marginal utility units,  $z_{t+1}$ .<sup>43</sup> Assume for example that the only shocks in the economy are fiscal shocks and that they take two values,  $g_H > g_L$ . We expect that the negative income effect of a fiscal shock reduces consumption and makes the household work more, leading to a smaller capital-labor ratio. As a result, we expect marginal utility and the marginal product of capital to increase when adverse fiscal shocks hit the economy. Thus, if the government hedges fiscal shocks by taking smaller positions against high shocks,  $z'_H < z'_L$ , the covariance will be *negative*, leading to an ex-ante capital *subsidy*.<sup>44</sup> Intuitively, the planner

---

<sup>42</sup>The difference in the two discount factors for the separable preferences (34) can be written as  $\frac{S_{t+1} - S_{t+1}^*}{S_{t+1}} = \frac{(1-\beta)(\gamma-\rho)\eta_{t+1}}{1/\Phi_{t+1}-\rho}$ . See the Appendix for details.

<sup>43</sup>A more complicated covariance criterion emerges when  $\gamma > \rho \neq 1$ :  $\bar{\tau}_{t+1}^K > (<) 0$  iff  $\text{Cov}_t^M(V_{t+1}^{\rho-1} \cdot U_{c,t+1} \cdot (1 - \delta + F_{K,t+1}), V_{t+1}^{\rho-1} z_{t+1}) > (<) 0$ .

<sup>44</sup>The covariance is  $\text{Cov}^M = E_t m_{t+1} c_{t+1}^{-1} (1 - \delta + F_{K,t+1}) \eta_{t+1}$ . Let subscripts denote if we are at the high or low shock and suppress time subscripts. By assumption we have  $c_H < c_L$ ,  $F_{K,H} > F_{K,L}$ ,  $\eta_H < 0$  and  $\eta_L > 0$ . Therefore,  $c_H^{-1}(1 - \delta + F_{K,H}) > c_L^{-1}(1 - \delta + F_{K,L})$ . The covariance is  $\text{Cov}^M = c_L^{-1}(1 - \delta + F_{K,L})\pi_L m_L \eta_L + c_H^{-1}(1 - \delta + F_{K,H})\pi_H m_H \eta_H$ . But  $c_L^{-1}(1 - \delta + F_{K,L})\eta_L < c_H^{-1}(1 - \delta + F_{K,H})\eta_L$ , since  $\eta_L > 0$ . Therefore,  $\text{Cov}^M <$

mitigates the effects of fiscal shocks by using a state-contingent capital subsidy at  $g_H$  and a state-contingent capital tax at  $g_L$ . But since adverse fiscal shocks are weighed more, we have an ex-ante subsidy to capital income. The Online Appendix provides a detailed example in an economy with a simplified stochastic structure (deterministic except for one period) that confirms this analysis.

## 9 Discussion: the case of $\rho > \gamma$

Consider now the case of  $\rho > \gamma$ . The direction of inequalities in propositions 2, 5 and 6 is obviously reversed.

**Proposition 8.** *(Desire to smooth over dates stronger than desire to smooth over states) Assume that  $\rho > \gamma$ , so that the household loves volatility in future utility. Then,  $\Phi_{t+1}(\hat{g}) < \Phi_t < \Phi_{t+1}(\tilde{g})$  when  $\hat{g}, \tilde{g}$  are such so that  $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$ . Similarly, in proposition 5 we have  $\tau_{t+1}(\hat{g}) < \tau_t < \tau_{t+1}(\tilde{g})$  when  $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$ . The same reversion of the direction of inequalities for  $\Phi_t$  holds also in an economy with capital, as in proposition 6. Proposition 7 goes through, but the direction of inequalities is reversed in (35):  $\bar{\tau}_{t+1}^K > (<) 0$  iff  $E_t \zeta_{t+1} \eta_{t+1} < (>) 0$ .*

Proposition 8 shows that the planner varies the excess burden over states and dates in the opposite way when  $\rho > \gamma$ . The underlying logic remains the same. Increases in continuation utility *increase* the stochastic discount factor when the household loves volatility in future utility (instead of decreasing it). Issuance of additional state-contingent debt reduces the stochastic discount factor, making debt relatively more expensive. Thus, the planner finds it optimal to “under-insure” in comparison to expected utility, selling less claims against good times and buying less claims against bad times. This is accompanied with smaller taxes in good times and higher taxes in bad times, amplifying the effects of fiscal shocks. Following the discussion in the previous section, there is an ex-ante capital *tax* instead of a subsidy, since bad times (which are weighed more) carry now a higher excess burden.

The martingale and submartingale results of propositions 3, 5 and 6 hold also for  $\rho > \gamma$ , so the persistence and back-loading results with respect to  $\pi_t \cdot M_t$  go through. The back-loading with respect to the physical measure goes through as well: the excess burden of taxation is now positively correlated with government spending. But the agent loves volatility in utility, and therefore places more probability mass on high-utility, low-spending shocks. Thus, we have again  $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) \leq 0$  and a positive drift with respect to the data-generating process.<sup>45</sup>

---

$c_H^{-1}(1 - \delta + F_{K,H})[\pi_L m_L \eta_L + \pi_H m_H \eta_H] = 0$ , since  $E m \eta = 0$ .

<sup>45</sup>In the Online Appendix I provide a full-blown quantitative exercise by setting  $\rho = 1$  and  $\gamma = 0$  for the constant Frisch utility function (26) with the same i.i.d. specification of shocks as in the baseline exercise. These are the RINCE preferences of Farmer (1990). The correlation of tax rates with government spending is highly positive and the autocorrelation of the tax rate close to unity, whereas the positive drift is small and discernible only in the long-run for this parametrization.

## 10 Concluding remarks

Dynamic fiscal policy revolves around the proper use of government debt returns in order to minimize the welfare loss of distortionary taxation. Consequently, empirically successful models of returns are a crucial ingredient in the determination of the optimal government debt portfolio and the resulting tax policy. Standard time-additive utility fails to capture basic asset-pricing facts, casting doubts on the conventional Ramsey policy prescriptions. This paper uses recursive preferences, a more promising preference specification for matching asset-pricing data, and re-evaluates the basic tenets of optimal fiscal policy.

I show how the tax-smoothing prescriptions of the dynamic Ramsey literature cease to hold with recursive utility. Optimal labor taxes become volatile and persistent. The planner mitigates the effects of fiscal shocks by taxing more in good times and less in bad times. Debt returns should be used to an even greater degree as a fiscal shock absorber, indicating that actual fiscal policy is even worse than we thought. Lastly, there is a novel incentive for the introduction of an intertemporal wedge, that can lead to ex-ante capital subsidies.

I have differentiated between time and risk attitudes in otherwise standard, complete markets economies of the dynamic Ramsey tradition. An analysis beyond the representative agent framework as in [Werning \(2007\)](#), [Bassetto \(2014\)](#) or [Bhandari et al. \(2015\)](#), or an exploration of different timing protocols like lack of commitment, are worthy directions for future research.

# A Economy without capital

## A.1 State space

At first, define

$$\begin{aligned}
A(g_1) &\equiv \left\{ (z_1, V_1) \mid \exists \{c_t, h_t\}_{t \geq 1}, \{z_{t+1}, V_{t+1}\}_{t \geq 1}, \text{ with } c_t \geq 0, h_t \in [0, 1] \right. \\
&\quad \text{such that:} \\
&\quad z_t = \Omega(c_t, h_t) + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}, t \geq 1 \\
&\quad V_t = \left[ (1-\beta)u(c_t, 1-h_t)^{1-\rho} + \beta \mu_t (V_{t+1})^{1-\rho} \right]^{\frac{1}{1-\rho}}, t \geq 1 \\
&\quad c_t + g_t = h_t, t \geq 1 \\
&\quad \text{where } m_{t+1} \text{ defined as in (5)} \\
&\quad \left. \text{and the transversality condition holds, } \lim_{t \rightarrow \infty} E_1 \beta^t \left( \frac{M_{t+1}}{M_1} \right)^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1} = 0. \right\}
\end{aligned}$$

The set  $A(g_1)$  stands for the set of values of  $z$  and  $V$  at  $t = 1$  that can be generated by an implementable allocation when the shock is  $g_1$ . From  $A(g)$  we get the state space as  $Z(g) \equiv \{z \mid \exists (z, V) \in A(g)\}$ .

## A.2 Transformed Bellman equation

Let  $v(z, g) \equiv \frac{V(z, g)^{1-\rho} - 1}{(1-\beta)(1-\rho)}$ . The Bellman equation takes the form

$$v(z, g) = \max_{c, h, z'_g} U(c, 1-h) + \beta \frac{\left[ \sum_{g'} \pi(g'|g) (1 + (1-\beta)(1-\rho)v(z'_g, g'))^{\frac{1-\gamma}{1-\rho}} \right]^{\frac{1-\rho}{1-\gamma}} - 1}{(1-\beta)(1-\rho)}$$

subject to the transformed implementability constraint

$$z = U_c c - U_h h + \beta \sum_{g'} \pi(g'|g) \frac{[1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{\rho-\gamma}{1-\rho}}}{[\sum_{g'} \pi(g'|g) [1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{1-\gamma}{1-\rho}}]^{\frac{\rho-\gamma}{1-\gamma}}} z'_g$$

and to (13)-(15). Recall that  $m'_{g'} \equiv \frac{V(z'_g, g')^{1-\gamma}}{\sum_{g'} \pi(g'|g) V(z'_g, g')^{1-\gamma}} = \frac{[1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{1-\gamma}{1-\rho}}}{\sum_{g'} \pi(g'|g) [1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{1-\gamma}{1-\rho}}}$ .

# B Economy with capital

## B.1 Competitive equilibrium

A price-taking firm operates the constant returns to scale technology. The firms rents capital and labor services and maximizes profits. Factor markets are competitive and therefore profit maximization leads to  $w_t = F_H(s^t)$  and  $r_t = F_K(s^t)$ .

The first-order condition with respect to an Arrow security is the same as in (9). The labor supply condition is  $U_l/U_c = (1 - \tau)w$ . The Euler equation for capital is

$$1 = \beta \sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) \left( \frac{V_{t+1}(s^{t+1})}{\mu_t(V_{t+1})} \right)^{\rho-\gamma} \frac{U_c(s^{t+1})}{U_c(s^t)} R_{t+1}^K(s^{t+1}). \quad (\text{B.1})$$

Conditions (9) and (B.1) deliver the no-arbitrage condition  $\sum_{s_{t+1}} p_t(s_{t+1}, s^t) R_{t+1}^K(s^{t+1}) = 1$ . The transversality conditions are

$$\lim_{t \rightarrow \infty} E_0 \beta^t M_t^{\frac{\rho-\gamma}{1-\gamma}} U_{ct} k_{t+1} = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} E_0 \beta^{t+1} M_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1} b_{t+1} = 0$$

## B.2 Ramsey problem

Define wealth as  $W_t(s^t) \equiv b_t(s^t) + R_t^K(s^t)k_t(s^{t-1})$ . Note that

$$\begin{aligned} \sum_{s_{t+1}} p_t(s_{t+1}, s^t) W_{t+1}(s^{t+1}) &= \sum_{s_{t+1}} p_t(s_{t+1}, s^t) [b_{t+1}(s^{t+1}) + R_{t+1}^K(s^{t+1})k_{t+1}(s^t)] \\ &= \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}) + k_{t+1}(s^t), \end{aligned}$$

by using the no-arbitrage condition. The household's budget constraint in terms of  $W_t$  becomes

$$c_t(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) W_{t+1}(s^{t+1}) = (1 - \tau_t(s^t)) w_t(s^t) h_t(s^t) + W_t(s^t).$$

Eliminate  $\{\tau_t, p_t\}$  and multiply with  $U_{ct}$  to get  $U_{ct} W_t = U_{ct} c_t - U_{lt} h_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1} W_{t+1}$ , which leads to the same implementability constraint for  $z_t \equiv U_{ct} W_t$ . At  $t = 0$  we have  $U_{c0} W_0 = U_{c0} c_0 - U_{l0} h_0 + \beta E_0 m_1^{\frac{\rho-\gamma}{1-\gamma}} z_1$ , where  $W_0 \equiv [(1 - \tau_0^K) F_K(s_0, k_0, h_0) + 1 - \delta] k_0 + b_0$ , and  $(k_0, b_0, \tau_0^K, s_0)$  given.

## B.3 Transformed Bellman equation with capital

Let  $v(z, k, s) \equiv \frac{V(z, k, s)^{1-\rho} - 1}{(1-\beta)(1-\rho)}$ . The Bellman equation takes the form

$$v(z, k, s) = \max_{c, h, k', z'_s} U(c, 1 - h) + \beta \frac{\left[ \sum_{s'} \pi(s'|s) (1 + (1 - \beta)(1 - \rho) v(z'_s, k', s')) \right]^{\frac{1-\rho}{1-\gamma}}}{(1 - \beta)(1 - \rho)} - 1$$

subject to

$$z = U_c c - U_l h + \beta \sum_{s'} \pi(s'|s) \frac{[1 + (1 - \beta)(1 - \rho)v(z'_{s'}, k', s')]^{\frac{\rho - \gamma}{1 - \rho}}}{[\sum_{s'} \pi(s'|s)[1 + (1 - \beta)(1 - \rho)v(z'_{s'}, k', s')]^{\frac{1 - \gamma}{1 - \rho}}]^{\frac{\rho - \gamma}{1 - \gamma}}} z'_{s'} \quad (\text{B.2})$$

$$c + k' - (1 - \delta)k + g_s = F(s, k, h) \quad (\text{B.3})$$

$$c, k' \geq 0, h \in [0, 1] \quad (\text{B.4})$$

The values  $(z'_{s'}, k')$  have to belong to the proper state space, i.e. it has to be possible that they can be generated by a competitive equilibrium with taxes that starts at  $(k, s)$ .

## B.4 First-order necessary conditions

$$c : U_c + \Phi \Omega_c = \lambda \quad (\text{B.5})$$

$$h : -U_l + \Phi \Omega_h = -\lambda F_H \quad (\text{B.6})$$

$$k' : \lambda = \beta \sum_{s'} \pi(s'|s) m'_{s'}{}^{\frac{\rho - \gamma}{1 - \gamma}} v_k(z'_{s'}, k', s') [1 + (1 - \beta)(\rho - \gamma)\eta'_{s'}\Phi] \quad (\text{B.7})$$

$$z'_{s'} : v_z(z'_{s'}, k', s') + \Phi [1 + (1 - \beta)(\rho - \gamma)v_z(z'_{s'}, k', s')\eta'_{s'}] = 0. \quad (\text{B.8})$$

$\Omega$  and  $\Omega_i, i = c, h$  are defined as in the proof of proposition 4. The relative *wealth* position  $\eta'_{s'}$  is defined as in (18) (with a value function  $V$  that also depends on capital now), so we again have  $\sum_{s'} \pi(s'|s) m'_{s'} \eta'_{s'} = 0$ . The envelope conditions are

$$v_z(z, k, s) = -\Phi \quad (\text{B.9})$$

$$v_k(z, k, s) = \lambda(1 - \delta + F_K). \quad (\text{B.10})$$

The envelope condition (B.9) together with (B.8) delivers the same law of motion of  $\Phi_t$  as in (19), leading to the same results as in proposition 3. Combine (B.5) and (B.6) and use the fact that  $(1 - \tau)F_H = U_l/U_c$  to get the same labor tax results as in propositions 4 and 5. Turn into sequence notation, use the law of motion of  $\Phi_t$  (19) to replace  $1 + (1 - \beta)(\rho - \gamma)\eta_{t+1}\Phi_t$  in (B.7) with the ratio  $\Phi_t/\Phi_{t+1}$  and the envelope condition (B.10) to eliminate  $v_k$  to finally get (32).

## B.5 Proof of proposition 7

The first-order condition with respect to consumption for  $t \geq 1$  is  $U_{ct} + \Phi_t \Omega_{ct} = \lambda_t$ . Thus,  $1/\Phi_t + \Omega_{ct}/U_{ct} = \lambda_t/(\Phi_t U_{ct}) > 0$ . Write the planner's discount factor as  $S_{t+1}^* = S_{t+1} \frac{\lambda_{t+1}/(\Phi_{t+1} U_{c,t+1})}{\lambda_t/(\Phi_t U_{ct})} = S_{t+1} \frac{1/\Phi_{t+1} + \Omega_{c,t+1}/U_{c,t+1}}{1/\Phi_t + \Omega_{ct}/U_{ct}}, t \geq 1$ . Remember that  $\Omega_c/U_c = 1 - \epsilon_{cc} - \epsilon_{ch}$ . Thus,

$$S_{t+1} - S_{t+1}^* = \frac{\frac{1}{\Phi_t} - \frac{1}{\Phi_{t+1}} + \epsilon_{cc,t+1} + \epsilon_{ch,t+1} - \epsilon_{cc,t} - \epsilon_{ch,t}}{\frac{1}{\Phi_t} + 1 - \epsilon_{cc,t} - \epsilon_{ch,t}} \cdot S_{t+1}, t \geq 1. \quad (\text{B.11})$$

The denominator is positive. Use (B.11) in the numerator of (33), simplify and normalize  $\zeta_{t+1}$  so that  $E_t \zeta_{t+1} = 1$  to get the criterion for capital taxation.

## References

- Ai, Hengjie and Ravi Bansal. 2016. Risk Preferences and The Macro Announcement Premium. Mimeo, Duke University.
- 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.
- Anderson, Evan W. 2005. The dynamics of risk-sensitive allocations. *Journal of Economic Theory* 125 (2):93–150.
- Bansal, Ravi and Amir Yaron. 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* LIX (4):1481–1509.
- Barro, Robert J. 1979. On the Determination of the Public Debt. *Journal of Political Economy* 87 (5):940–71.
- Bassetto, Marco. 2014. Optimal Fiscal Policy with Heterogenous Agents. *Quantitative Economics* 5 (3):675–704.
- Berndt, Antje, Hanno Lustig, and Sevin Yeltekin. 2012. How Does the U.S. Government Finance Fiscal Shocks? *American Economic Journal: Macroeconomics* 4:69–104.
- Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent. 2015. Taxes, debts, and redistributions with aggregate shocks. Mimeo, New York University.
- . 2016. Fiscal Policy and Debt Management with Incomplete Markets. *Quarterly Journal of Economics* Forthcoming.
- Campbell, John Y. 1993. Intertermoral asset pricing without consumption data. *American Economic Review* 83 (3):487–512.
- Chamley, Christophe. 1986. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54 (3):607–622.
- Chari, V.V. and Patrick J. Kehoe. 1999. Optimal fiscal and monetary policy. *Handbook of Macroeconomics* 1:1671–1745.
- 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.
- Croce, Mariano M., Howard Kung, Thien T. Nguyen, and Lukas Schmid. 2012a. Fiscal Policies and Asset Prices. *The Review of Financial Studies* 25 (9):2635–2672.
- Croce, Mariano M., Thien T. Nguyen, and Lukas Schmid. 2012b. The market price of fiscal uncertainty. *Journal of Monetary Economics* 59 (5):401–416.
- Epstein, Larry G. and Stanley E. Zin. 1989. Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57 (4):937–969.
- Epstein, Larry G., Emmanuel Farhi, and Tomasz Strzalecki. 2014. How Much Would You Pay to Resolve Long-Run Risk? *American Economic Review* 104 (9):2680–97.

- Farhi, Emmanuel. 2010. Capital Taxation and Ownership when Markets are Incomplete. *Journal of Political Economy* 118 (5):908–948.
- Farhi, Emmanuel and Iván Werning. 2008. Optimal savings distortions with recursive preferences. *Journal of Monetary Economics* 55 (1):21–42.
- Farmer, Roger E.A. 1990. Rince Preferences. *The Quarterly Journal of Economics* 105 (1):43–60.
- Gomes, Francisco, Alexander Michaelides, and Valery Polkovnichenko. 2013. Fiscal Policy and Asset Prices with Incomplete Markets. *The Review of Financial Studies* 26 (2):531–566.
- Gottardi, Piero, Atsushi Kajii, and Tomoyuki Nakajima. 2015. Optimal Taxation and Debt with Uninsurable Risks to Human Capital Accumulation. *American Economic Review* 105 (11):3443–70.
- Gourinchas, Pierre-Olivier and Hélène Rey. 2007. International Financial Adjustment. *Journal of Political Economy* 115 (4):665–703.
- Gourio, François. 2012. Disaster Risk and Business Cycles. *American Economic Review* 102 (6):2734–66.
- Hall, George J. and Stephan Krieger. 2000. The Tax Smoothing Implications of the Federal Debt Paydown. *Brookings Papers on Economic Activity* 31:253–302.
- Hall, George J. and Thomas J. Sargent. 2011. Interest Rate Risk and Other Determinants of Post-WWII US Government Debt/GDP Dynamics. *American Economic Journal: Macroeconomics* 3 (3):192–214.
- Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.
- Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008. Consumption Strikes Back? Measuring Long-Run Risk. *Journal of Political Economy* 116 (2):260–302.
- Judd, Kenneth L. 1985. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28 (1):59–83.
- Karantounias, Anastasios G. 2013. Managing pessimistic expectations and fiscal policy. *Theoretical Economics* 8 (1):193–231.
- Koopmans, Tjalling C. 1960. Stationary ordinary utility and impatience. *Econometrica* 28 (2):287–309.
- Kreps, D. M. and E. L. Porteus. 1978. Temporal Resolution of Uncertainty and Dynamic Choice. *Econometrica* 46:185–200.
- Kydland, Finn E. and Edward C. Prescott. 1980. Dynamic optimal taxation, rational expectations and optimal control. *Journal of Economic Dynamics and Control* 2:79–91.
- Lucas, Robert Jr. and Nancy L. Stokey. 1983. Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics* 12 (1):55–93.

- Marcet, Albert and Andrew Scott. 2009. Debt and deficit fluctuations and the structure of bond markets. *Journal of Economic Theory* 144 (2):473–501.
- Petrosky-Nadeau, Nicolas, Lu Zhang, and Lars-Alexander Kuehn. 2013. Endogenous disasters and asset prices. Mimeo, Carnegie Mellon University.
- Piazzesi, Monika and Martin Schneider. 2007. Equilibrium Yield Curves. *NBER Macroeconomics Annual* 389–442. Edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford.
- Rudebusch, Glenn D. and Eric T. Swanson. 2012. The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. *American Economic Journal: Macroeconomics* 4 (1):105–143.
- Shin, Yongseok. 2006. Ramsey meets Bewley: Optimal Government Financing with Incomplete Markets. Mimeo, Washington University in St. Louis.
- Sleet, Christopher. 2004. Optimal taxation with private government information. *Review of Economic Studies* 71 (4):1217–1239.
- Sleet, Christopher and Sevin Yeltekin. 2006. Optimal taxation with endogenously incomplete debt markets. *Journal of Economic Theory* 127 (1):36–73.
- Strzalecki, Tomasz. 2011. Axiomatic Foundations of Multiplier Preferences. *Econometrica* 79 (1):47–73.
- . 2013. Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion. *Econometrica* 81 (3):1039–1074.
- Tallarini, Thomas D. Jr. 2000. Risk-sensitive real business cycles. *Journal of Monetary Economics* 45 (3):507–532.
- Weil, Philippe. 1990. Non-Expected Utility in Macroeconomics. *Quarterly Journal of Economics* CV (1):29–42.
- Werning, Iván. 2007. Optimal fiscal policy with redistribution. *Quarterly Journal of Economics* 122 (3):925–967.
- Zhu, Xiaodong. 1992. Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58 (2):250–289.

# Optimal fiscal policy with recursive preferences

Anastasios G. Karantounias \*

October 23, 2017

## Online Appendix (not for publication)

JEL CLASSIFICATION: D80; E62; H21; H63.

KEY WORDS: Ramsey plan, tax smoothing, Epstein-Zin, recursive utility, excess burden, labor tax, capital tax, martingale, fiscal insurance.

---

\*Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St NE, Atlanta, GA 30309, [anastasios.karantounias@atl.frb.org](mailto:anastasios.karantounias@atl.frb.org). All errors are my own. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

# Contents

<b>A</b>	<b>Initial period problem</b>	<b>3</b>
A.1	Economy without capital . . . . .	3
A.2	Economy with capital . . . . .	3
A.3	Optimality conditions at $t = 0$ . . . . .	4
<b>B</b>	<b>Martingales and (non)-convergence</b>	<b>4</b>
<b>C</b>	<b>Computational details</b>	<b>5</b>
C.1	Solution method . . . . .	5
C.2	Size of the state space . . . . .	7
C.2.1	Upper bounds for $z$ and the surplus in marginal utility units . . . . .	7
C.2.2	Upper bounds and candidate convergence points . . . . .	10
C.2.3	Non-convexities and robustness checks for the baseline exercise . . . . .	13
<b>D</b>	<b>Higher risk aversion or more volatile shocks</b>	<b>18</b>
<b>E</b>	<b>Persistent shocks and fiscal insurance exercise</b>	<b>20</b>
E.1	Calibration of shocks, policy functions and some sample paths . . . . .	20
E.2	Moments of interest . . . . .	23
E.3	Returns and risk premia . . . . .	25
E.4	Fiscal insurance in <a href="#">Berndt et al. (2012)</a> and log-linear approximation constants . . . . .	29
E.5	News variables . . . . .	30
E.6	Linear approximation . . . . .	32
E.7	Effect of approximation constants . . . . .	35
<b>F</b>	<b>Example in an economy with capital</b>	<b>36</b>
F.1	Analysis . . . . .	36
F.2	Computational details . . . . .	38
<b>G</b>	<b>Sequential formulation of Ramsey problem</b>	<b>41</b>
<b>H</b>	<b>Preference for late resolution of uncertainty</b>	<b>44</b>

# A Initial period problem

## A.1 Economy without capital

The problem at  $t = 0$  is

$$\bar{V}_0(b_0, g_0) \equiv \max_{c_0, h_0, z_{1, g_1}} \left[ (1 - \beta)u(c_0, 1 - h_0)^{1-\rho} + \beta \left[ \sum_{g_1} \pi_1(g_1|g_0) V(z_{1, g_1}, g_1)^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

subject to

$$U_{c_0} b_0 = U_c c_0 - U_l h_0 + \beta \sum_{g_1} \pi_1(g_1|g_0) \frac{V(z_{1, g_1}, g_1)^{\rho-\gamma}}{\left[ \sum_{g_1} \pi_1(g_1|g_0) V(z_{1, g_1}, g_1)^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z_{1, g_1} \quad (\text{A.1})$$

$$c_0 + g_0 = h_0 \quad (\text{A.2})$$

$$c_0 \geq 0, h_0 \in [0, 1], \quad (\text{A.3})$$

$$z_{1, g_1} \in Z(g_1) \quad (\text{A.4})$$

where  $(b_0, g_0)$  given. The notation  $z_{1, g_1}$  denotes the value of the state variable  $z_1$  at  $g_1$ . The overall value of the Ramsey problem  $\bar{V}(\cdot)$  and the initial period policy functions  $(c_0, h_0, z_1)$  depend on the initial conditions  $(b_0, g_0)$ .

## A.2 Economy with capital

The problem at  $t = 0$  in an economy with capital is

$$\bar{V}_0(b_0, k_0, s_0, \tau_0^K) \equiv \max_{c_0, h_0, k_1, z_{1, s_1}} \left[ (1 - \beta)u(c_0, 1 - h_0)^{1-\rho} + \beta \left[ \sum_{s_1} \pi_1(s_1|s_0) V(z_{1, s_1}, k_1, s_1)^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

subject to

$$U_{c_0} \left[ ((1 - \tau_0^K) F_K(s_0, k_0, h_0) + 1 - \delta) k_0 + b_0 \right] = U_c c_0 - U_l h_0 + \beta \sum_{s_1} \pi_1(s_1|s_0) \frac{V(z_{1, s_1}, k_1, s_1)^{\rho-\gamma}}{\left[ \sum_{s_1} \pi_1(s_1|s_0) V(z_{1, s_1}, k_1, s_1)^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z_{1, s_1} \quad (\text{A.4})$$

$$c_0 + k_1 - (1 - \delta) k_0 + g_0 = F(s_0, k_0, h_0) \quad (\text{A.5})$$

$$c_0, k_1 \geq 0, h_0 \in [0, 1], \quad (\text{A.6})$$

where  $(b_0, k_0, s_0, \tau_0^K)$  given. Same comments apply for the dependence on the initial conditions.

### A.3 Optimality conditions at $t = 0$

Consider the economy with capital. Let  $\Phi_0$  and  $\lambda_0$  denote the respective multipliers on the initial period implementability and resource constraint of the transformed problem and recall the definition of  $\Omega$  in the text. The initial period optimality conditions are:

$$c_0 : \quad U_{c0} + \Phi_0 \left[ \Omega_{c0} - U_{cc,0} W_0 \right] = \lambda_0 \quad (\text{A.7})$$

$$h_0 : \quad -U_{l0} + \Phi_0 \left[ \Omega_{h0} + U_{cl,0} W_0 - U_{c0} (1 - \tau_0^K) F_{KH,0} k_0 \right] = -\lambda_0 F_{H0} \quad (\text{A.8})$$

$$k_1 : \quad \lambda_0 = \beta \sum_{s_1} \pi(s_1|s_0) m_{1,s_1}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z_{1,s_1}, k_1, s_1) \left[ 1 + (1 - \beta)(\rho - \gamma) \eta_{1,s_1} \Phi_0 \right] \quad (\text{A.9})$$

$$\begin{aligned} z_{1,s_1} : \quad & v_z(z_{1,s_1}, k_1, s_1) \\ & + \Phi_0 \left[ 1 + (1 - \beta)(\rho - \gamma) v_z(z_{1,s_1}, k_1, s_1) \eta_{1,s_1} \right] = 0, \end{aligned} \quad (\text{A.10})$$

where  $W_0 = [(1 - \tau_0^K) F_K(s_0, k_0, h_0) + (1 - \delta)] k_0 + b_0$ , the household's initial wealth, and  $\eta_{1,s_1}$  the household's relative wealth position in marginal utility units at  $s_1$ . The initial period first-order conditions for an economy without capital for the variables  $(c_0, h_0, z_{1,g_1})$ , are (A.7), (A.8) and (A.10) with  $W_0 = b_0$ ,  $F_{H0} = 1$ ,  $F_{KH} \equiv 0$ .

Using the same steps as in the labor tax proposition we get the optimal labor tax at  $t = 0$ ,

$$\tau_0 = \Phi_0 \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc} + (1 - \tau_0^K) \frac{F_{KH}}{F_H} k_0 - (\epsilon_{cc} + \epsilon_{hc}) c_0^{-1} W_0}{1 + \Phi_0 (1 + \epsilon_{hh} + \epsilon_{hc} - \epsilon_{hc} c_0^{-1} W_0)},$$

which simplifies to

$$\tau_0 = \Phi_0 \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc} - (\epsilon_{cc} + \epsilon_{hc}) c_0^{-1} b_0}{1 + \Phi_0 (1 + \epsilon_{hh} + \epsilon_{hc} - \epsilon_{hc} c_0^{-1} b_0)},$$

in an economy without capital. All elasticities in the two formulas are evaluated at the  $t = 0$  allocation.

## B Martingales and (non)-convergence

Note that the ratio  $M_t/\Phi_t$  is a martingale with respect to  $\pi$ :  $E_t(M_{t+1}/\Phi_{t+1}) = M_t E_t m_{t+1} (1/\Phi_{t+1}) = M_t/\Phi_t$ , since  $1/\Phi_t$  is a martingale with respect to  $\pi \cdot M$ . Therefore, by the martingale convergence theorem, the non-negative ratio  $M_t/\Phi_t$  converges almost surely with respect to  $\pi$  to a finite random variable. Furthermore, since  $M_t$  is by construction a non-negative martingale with respect to  $\pi$ , it converges to the non-negative random variable  $M_\infty$  almost surely. Thus, we know that the product of  $M_t$  and  $1/\Phi_t$  converges and that  $M_t$  converges. However, we cannot make a general claim about the convergence of  $1/\Phi_t$ , unless we restrict the analysis to the case of an absorbing

state. Why? If  $M_\infty > 0$ , then we could infer that  $1/\Phi_t$  converges almost surely with respect to  $\pi$ . Alas, the martingale  $M_t$  typically converges to zero, so we cannot make this claim.

To understand better the issue of  $M_\infty = 0$ , we can follow similar steps as [Aiyagari et al. \(2002\)](#) (who worked with the risk-adjusted measure) and show that if  $M_\infty(\tilde{\omega}) > 0$  for a sample path  $\tilde{\omega}$ , then the increment has to converge to unity,  $m_t(\tilde{\omega}) \rightarrow 1$ , so  $V_t^{1-\gamma}(\tilde{\omega}) \rightarrow E_{t-1}V_t^{1-\gamma}(\tilde{\omega})$ . The logic is simple:  $\ln M_t(\tilde{\omega}) = \sum_{i=1}^t \ln m_i(\tilde{\omega}) \rightarrow \ln M_\infty(\tilde{\omega}) > -\infty$  (since  $M_\infty(\tilde{\omega}) > 0$ ) and therefore  $\ln m_t(\tilde{\omega}) \rightarrow 0$ . Thus, we can infer that if  $Pr(\tilde{\omega} | m_t(\tilde{\omega}) \rightarrow 1) = 0$ , then  $M_\infty = 0$  almost surely (otherwise  $m_t \rightarrow 1$  on a set of positive measure). Actually, this result can be strengthened to the following statement: if it is *not* the case that  $m_t \rightarrow 1$  almost surely, then  $M_\infty = 0$  almost surely. The proof of this is coming from the work of Ian Martin who generalized the Kakutani theorem on multiplicative martingales. See [Martin \(2012, Theorem 1\)](#). Thus, as long as there is some positive probability that there is variation in continuation values at the limit so that  $m_{t+1} \not\rightarrow 1$ , we run into the case of  $M_\infty = 0$ .<sup>1</sup>

To conclude, the martingale property of the inverse of the excess burden of taxation does not provide sufficient information for establishing convergence results with respect to  $\pi$ . Additional information (like properties of the utility function etc) and a careful numerical analysis is needed.

## C Computational details

### C.1 Solution method

**State space.** Combine the first-order conditions with respect to consumption and leisure to get the optimal wedge in labor supply as  $\frac{U_l}{U_c} \cdot \frac{1-\Phi \frac{\Omega_h}{U_l}}{1+\Phi \frac{\Omega_c}{U_c}} = 1$ . Using the optimal wedge and the resource constraint, we can express the optimal consumption-labor allocation as functions of the shock  $g$  and  $\Phi$ ,  $c(g, \Phi)$  and  $h(g, \Phi)$ . Let  $U^*(g, \Phi) \equiv U(c(g, \Phi), 1 - h(g, \Phi))$  and  $\Omega^*(g, \Phi) \equiv \Omega(c(g, \Phi), h(g, \Phi))$ .  $U^*$  stands for the period utility at  $g$  when the excess burden of taxation is  $\Phi$  and  $\Omega^*$  the respective government surplus in marginal utility units. For the utility function in the baseline exercise we have  $\Omega^* = 1 - a_h h(g, \Phi)^{1+\phi_h}$ .

To generate values of  $z$ , fix  $\Phi$  to a particular value. Given a constant value of  $\Phi$  we get a history-independent allocation which allows us to solve easily for the the utility recursion  $v^*(g, \Phi) = U^*(g, \Phi) + \frac{\beta}{(1-\beta)(1-\gamma)} \ln \sum_{g'} \pi(g'|g) \exp((1-\beta)(1-\gamma)v^*(g', \Phi))$ . For each given  $\Phi$  we get also the induced conditional likelihood ratio  $m(g'|g) = \exp((1-\beta)(1-\gamma)v^*(g', \Phi)) / \sum_{g'} \pi(g'|g) \exp((1-\beta)(1-\gamma)v^*(g', \Phi))$ . The induced debt positions  $z$  for a given  $\Phi$  are

$$\mathbf{z} = (I - \beta \tilde{\Pi})^{-1} \mathbf{\Omega}^*,$$

---

<sup>1</sup>In the quantitative exercise the likelihood ratio  $M_t$  does indeed converge to zero. In [Aiyagari et al. \(2002\)](#) we typically have convergence of the risk-adjusted measure to zero.

where boldface variables denote column vectors and  $\tilde{\Pi} \equiv \Pi \circ \mathcal{M}$ , where  $\Pi$  the transition matrix of the shocks and  $\mathcal{M}$  the matrix of  $m(g'|g)$ . The symbol  $\circ$  denotes element by element (or else Hadamard) multiplication.

The induced values of  $z$  can be generated – by construction – in a competitive equilibrium and are a “nice” subset of the true state space. I vary  $\Phi$  in the set  $[0, \bar{\Phi}]$ . The zero value of  $\Phi$  corresponds to the first-best allocation, so the induced  $z$ 's are the level of government *assets* necessary to finance government expenditures without having to resort to distortionary taxation. I will talk further about the choice of  $\bar{\Phi}$  in the next section. Let  $Z_i$  denote the state space for the low and high shock,  $i = L, H$ . For the lower and upper bounds of  $Z_i$  I use the minimum and maximum value of the debt position at  $i$  generated by a  $\Phi$  in  $[0, \bar{\Phi}]$  (which just correspond to  $\Phi = 0$  and  $\Phi = \bar{\Phi}$ , because the implied  $z$ 's are an increasing function of  $\Phi$ ).

**Initial estimate of the value function.** For each  $\Phi$ , I can associate the induced  $z$  to an induced  $v^*$ , which provides an initial guess for the value function,  $v^0(z, g_i)$ ,  $z \in Z_i$ . At first, I form a grid of points for  $Z_i, i = L, H$  and perform value function iteration with grid search. There may be convergence issues because updating the value function in the constraint hinders the contraction property. To avoid that I have two loops:

- Inner Loop: *Given* the value function in the constraint, iterate on the Bellman equation till convergence (I use also policy function iteration to increase speed).
- Outer Loop: Update the value function in the constraint and repeat the inner loop.

The procedure is stopped when the value function in the constraint is approximately equal to the value function in the Bellman equation. The inner loop entails standard value function iteration and is convergent. There is no guarantee of convergence of the double loop. In the outer loop I use damping in order to improve convergence properties.

**Final estimate of the value function.** I use grid search (with the constant- $\Phi$  first guess  $v^*$ ) in order to avoid non-convexity issues and the possibility of a local optimum. This procedure provides a first estimate of the value functions. For improved *precision*, I use the output of the double-loop procedure as an initial guess and fit the value functions at the two shocks with cubic splines. At the final stage the value function in the constraint is updated every period. I use 167 breakpoints and 500 points for each  $Z_i$  and apply regression. More grid points are allocated at the upper half of each state space in order to capture better the curvature of the value functions. A continuous optimization routine is used, with initial guesses the policy functions that came from the grid search.

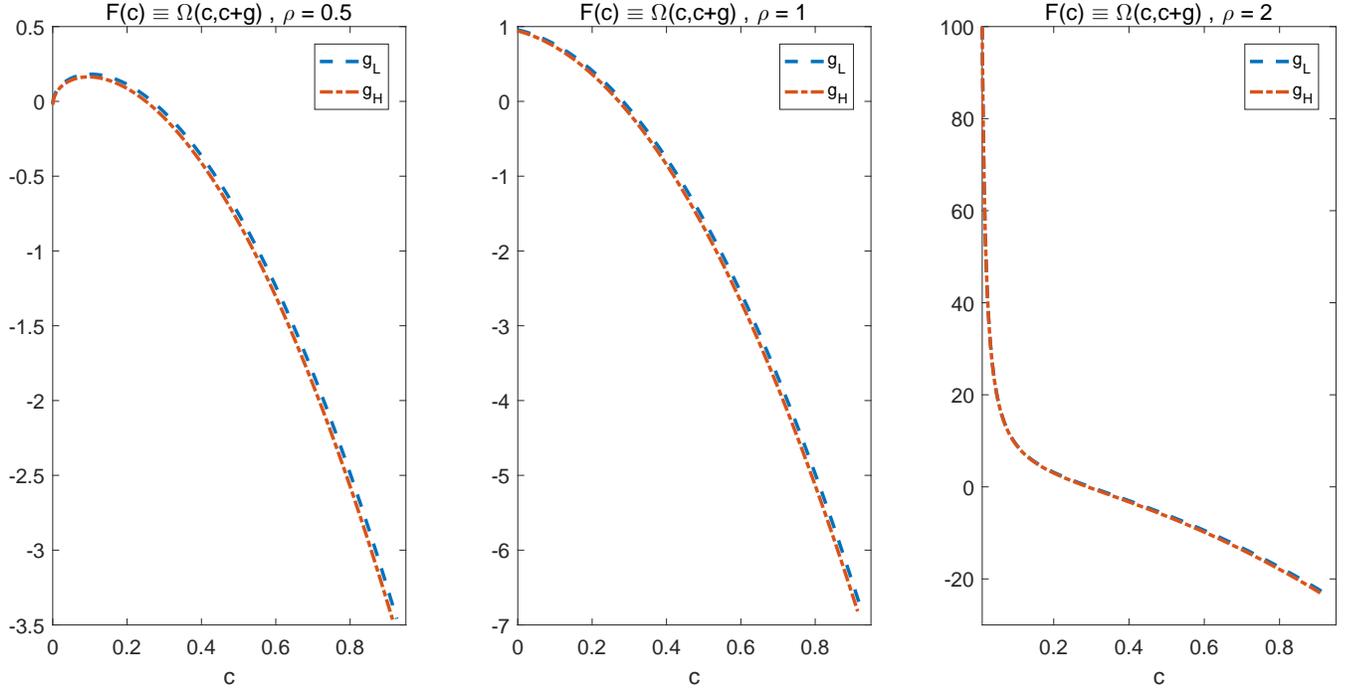


Figure C.1: The graphs plot from left to right the equilibrium surplus in marginal utility units as function of consumption  $c$  for  $\rho = 0.5, 1, 2$ . For  $\rho = 0.5$  the maximizer is interior. That point corresponds to  $\tau \rightarrow \hat{\tau}$  and  $\Phi \rightarrow \infty$ . For  $\rho = 1$ , the upper bound corresponds to  $c = 0$  and  $\tau \rightarrow 1, \Phi \rightarrow \infty$ . For  $\rho > 1$  there is no upper bound in the surplus in marginal utility units when consumption approaches zero.

## C.2 Size of the state space

There are several considerations about the size of the state space and the choice of the upper bound  $\bar{\Phi}$ , that we now turn to.

### C.2.1 Upper bounds for $z$ and the surplus in marginal utility units

To understand the maximum amount of debt in marginal utility units, we have first to understand the behavior of the surplus in marginal utility units. If this is unbounded, then the present discounted value of surpluses (where the discounting involves also continuation utilities coming from the richer pricing kernels) will be unbounded. Define  $F(c) \equiv \Omega(c, c+g)$ , that is, the surplus in marginal utility units as function of equilibrium consumption. Assume for simplicity that  $U_{cl} \geq 0$ . By recalling the elasticity formulas that were used in the proof of the labor tax proposition 4 and using the fact that  $1 - \tau = U_l/U_c$ , we have

$$F'(c) = \Omega_c + \Omega_h = U_c [1 - \epsilon_{cc} - \epsilon_{ch} - (1 - \tau)(1 + \epsilon_{hh} + \epsilon_{hc})]$$

Then,

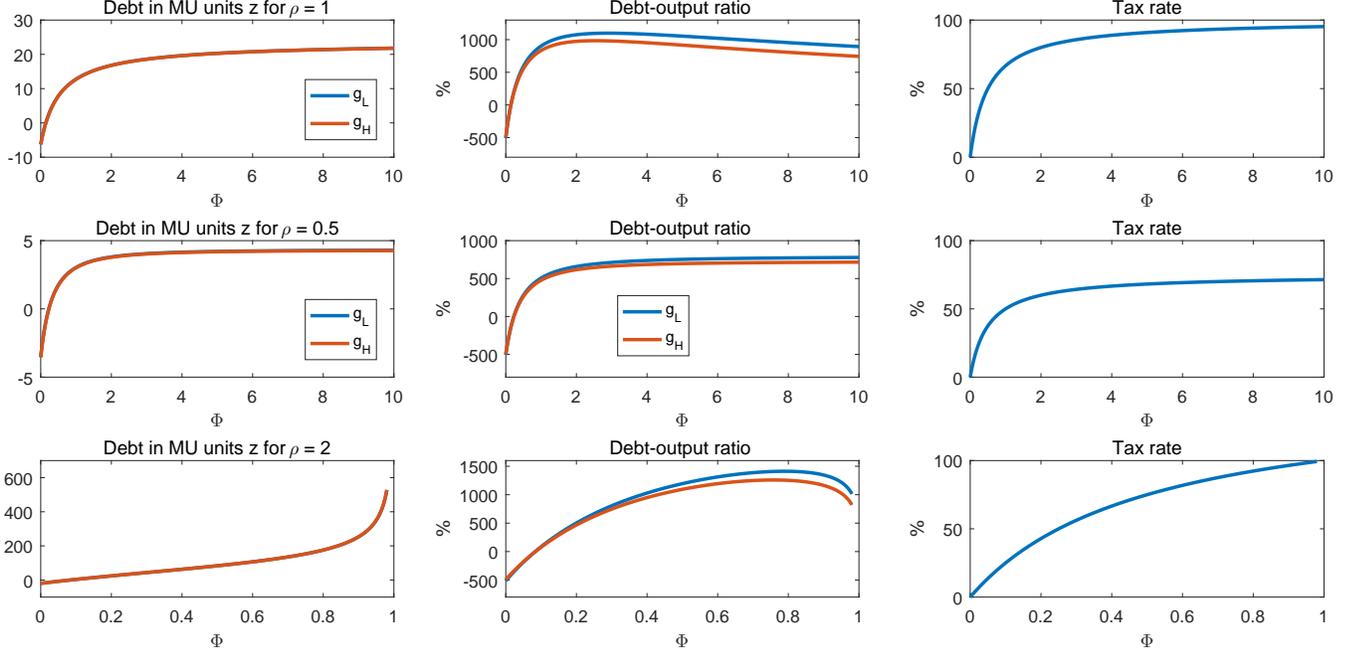


Figure C.2: The panels depict  $z$ , the corresponding debt ratio and the tax rate that are induced by a constant- $\Phi$  policy. The top row corresponds to the baseline exercise of  $\rho = 1$ . The middle row corresponds to  $\rho = 0.5$  and the bottom row to  $\rho = 2$ . For  $\rho \neq 1$ , I have considered a risk-sensitive recursion. Recall that the maximum  $\Phi$  is infinity for the first two rows with a respective tax rate of  $\hat{\tau}$  and the upper bounds of  $z$  that are not attained. For the bottom row, the maximum excess burden is  $\Phi = 1/(\rho - 1) = 1$  and  $z$  is unbounded.

$$F'(c) < 0 \Rightarrow \tau < \hat{\tau} \equiv \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc}}{1 + \epsilon_{hh} + \epsilon_{hc}}$$

But we know from proposition 4 that the optimal labor tax is  $\tau = \frac{\Phi(\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc})}{1 + \Phi(\epsilon_{hh} + \epsilon_{hc})}$ , so  $\tau < \hat{\tau}$  for any finite  $\Phi$ . Thus, the tax rate would never be so high (and the consumption so low) so that  $F'(c) > 0$  (the discussion is obviously about the correct side of the Laffer curve in terms of the surplus in marginal utility units). Obviously, the value of  $F$  at the level of consumption that corresponds to  $\hat{\tau}$ , *if* finite, is a candidate for an upper bound of the surplus in marginal utility units.

Consider now the period utility function of proposition 5, i.e. power utility with constant Frisch elasticity. We have  $\hat{\tau} = \frac{\rho + \phi_h}{1 + \phi_h}$ . Before analyzing the behavior of  $F(c)$ , let's examine the excess burden of taxation and the implied tax rate. From the first-order condition with respect to consumption in the Ramsey problem we have  $c^{-\rho}(1 + \Phi(1 - \rho)) = \lambda > 0$  (time subscripts are omitted). The positivity of the multiplier  $\lambda$  implies a positivity restriction on  $1 + \Phi(1 - \rho) > 0$ . This restriction is not binding for  $\rho \leq 1$ , so the excess burden of taxation could in principle diverge to infinity. For  $\rho > 1$ , this restriction imposes an upper bound for  $\Phi$ ,  $\Phi < \frac{1}{\rho - 1}$ . Consider also the

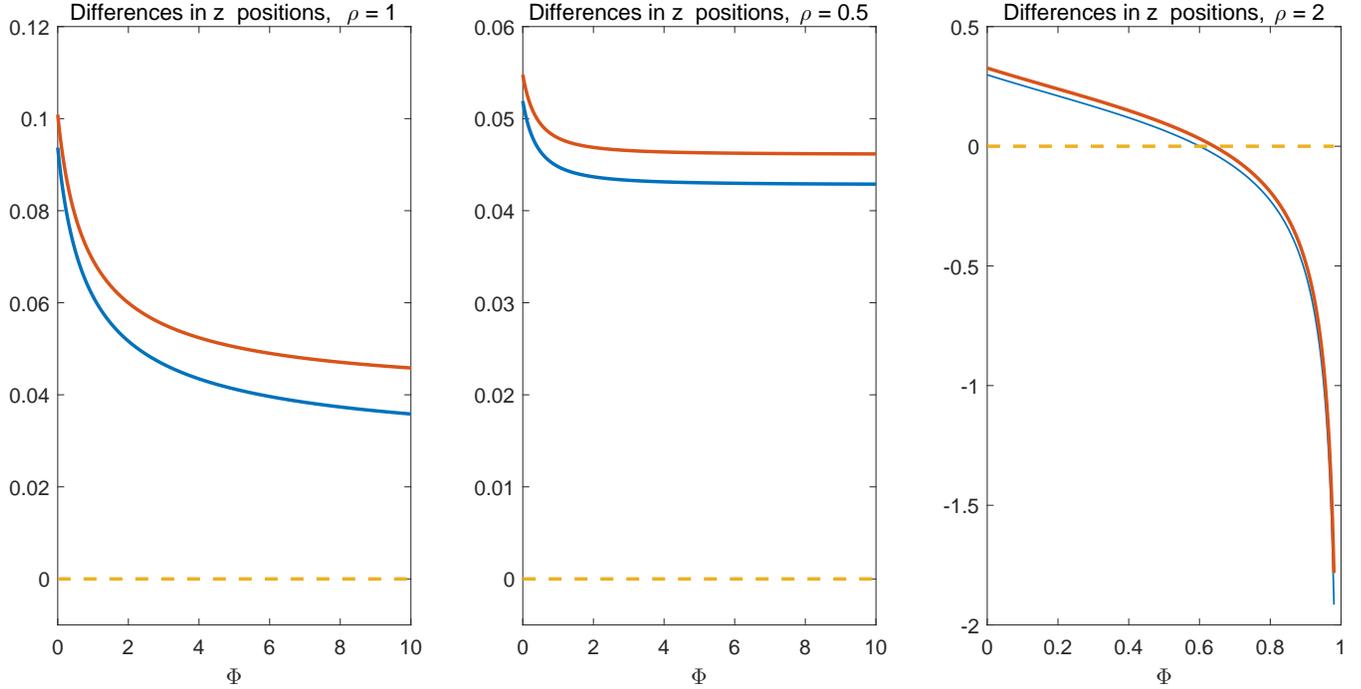


Figure C.3: Positions  $z$  induced by a constant- $\Phi$  policy for different  $\rho$ . I use a risk-sensitive recursion for  $\rho \neq 1$ . The shock takes three values: low (L), medium (M) and high (H). The differences  $z_L - z_M$  (blue line) and  $z_M - z_H$  (red line) are depicted.

tax rate,  $\tau(\Phi) = \frac{\Phi(\rho+\phi_h)}{1+\Phi(1+\phi_h)}$ . For  $\rho < 1$  we have  $\lim_{\Phi \rightarrow \infty} \tau(\Phi) = \frac{\rho+\phi_h}{1+\phi_h} = \hat{\tau} < 1$ . For  $\rho = 1$ , we have  $\lim_{\Phi \rightarrow \infty} \tau(\Phi) = \hat{\tau} = 1$ . For  $\rho > 1$ , we have  $\lim_{\Phi \rightarrow 1/(\rho-1)} \tau(\Phi) = 1 < \hat{\tau}$ .

Turn now to the surplus in marginal utility units. We have  $F(c) = c^{1-\rho} - a_h(c+g)^{1+\phi_h}$ , with  $F'(c) = c^{-\rho}[1 - \rho - (1 - \tau)(1 + \phi_h)]$ . Obviously, for  $\rho < 1$  we could have  $F'(c) > 0$  and the consumption  $\hat{c}$  corresponding to the upper bound is interior. The respective excess burden approaches infinity and the tax rate approaches  $\hat{\tau}$ . For  $\rho = 1$ , the upper bound corresponds to  $c = 0$ , taking the value  $F(0) = 1 - a_h g^{1+\phi_h}$ . This corresponds to a tax rate that becomes asymptotically 100% (and  $\Phi \rightarrow \infty$ ). For  $\rho > 1$  there is no upper bound, since the surplus in marginal utility units approaches infinity when  $c \rightarrow 0$  (and  $\tau \rightarrow 1$ ). Clearly, even when there is an upper bound, it will not be attained. Figure C.1 depicts the three cases.

Figure C.2 considers the induced  $z$  by a constant- $\Phi$  policy, together with the respective debt ratios and tax rates. The top row considers the baseline exercise for  $\rho = 1$  (where there is an upper bound that corresponds to a tax rate converging to 100%). The middle and bottom row consider the cases of  $\rho = 0.5$  (where there is an upper bound), and the unbounded case of  $\rho = 2$  respectively. Without loss of generality (concerning the discussion about bounds), I have considered risk-sensitive preferences for the cases of  $\rho \neq 1$  (by setting  $\sigma = (1 - \beta)(1 - \gamma)$ ).

### C.2.2 Upper bounds and candidate convergence points

The previous section showed that there may not be a natural upper bound for  $z$ , or if there is, it may not be attained. For any type of computation though we need to make a choice for  $\bar{\Phi}$ . An issue of concern may be that there is a positive convergence point for the excess burden that cannot be reached because the upper bound we chose is not large enough. The following proposition shows that this is not the case for the baseline exercise with logarithmic utility,  $\rho = 1$ .

**Proposition C.1.** *Consider the utility function of the baseline exercise and the i.i.d. shock specification. If the excess burden of taxation does converge, then it necessarily has to converge to zero.*

*Proof.* Assume that  $\Phi_t$  converges along a sample path to the value  $\Phi$  (which may depend on the sample path). Recall that  $\Omega^*(g, \Phi) = 1 - a_h h(g, \Phi)^{1+\phi_h}$ . Use the implicit function theorem in the two-equation system formed by the optimal wedge equation and the resource constraint to get  $\partial h/\partial g = h/(h + \phi_h c) > 0$  and  $\partial c/\partial g = -\phi_h c/(h + \phi_h c) < 0$ . Thus,  $\partial \Omega^*/\partial g = -a_h(1 + \phi_h)h^{\phi_h} \partial h/\partial g < 0$ . Therefore, the surplus in marginal utility units is always *larger* for the smaller shock for *any* value of  $\Phi$ . As a result, debt in marginal utility units is always higher for the lower shock, since for a constant  $\Phi$  we have  $z(g, \Phi) = \Omega^*(g, \Phi) + \frac{\beta}{1-\beta} \sum_{g'} \pi(g') m(g') \Omega^*(g', \Phi)$  ( $m(g')$  stand for the conditional likelihood ratio induced by the constant  $\Phi$ . It does not depend on the current  $g$  due to the i.i.d. assumption). But then for any  $\Phi > 0$  the planner will always increase the excess burden of taxation for low shocks, since  $\Phi'_{g'} = \Phi/(1 + (1-\beta)(1-\gamma)\eta'_{g'}\Phi)$ , contradicting the premise of a constant  $\Phi$ . Only in the case of a *zero*  $\eta'_{g'}$ ,  $\forall g'$ , i.e. only if there was a  $\Phi > 0$  such that  $z$  is equal across shocks, would it be possible to have a constant  $\Phi$ . This cannot be the case, as proved earlier. The only option of having a constant  $\Phi$  would be to have  $\Phi = 0$ , which implies that the second-best allocation converges to the first-best. In that case, the first-best is an absorbing state, and the government is using the interest income on accumulated assets to finance government spending. Note that the i.i.d. assumption in the proposition was used only to guarantee that debt in marginal utility units varies across shocks as  $\Omega^*$  does. Persistent shocks could also be allowed as long as the implied  $z$ 's do vary across shocks.  $\square$

The heart of the proof in proposition C.1 is the following: at any positive candidate convergence point  $\Phi$ , the induced  $z$  policies have to be the *same* across shocks, because otherwise they would imply a non-zero relative debt position  $\eta_{t+1}$ . A non-zero relative debt position would make the planner change the excess burden of taxation across states of the world. The proof shows that the induced  $z$  positions *do* vary across shocks for any given  $\Phi$  when  $\rho = 1$ . The left panel in figure C.3 confirms that. It plots the differences in  $z$ ,  $z_L - z_M$  and  $z_M - z_H$ , when the shock takes three values: low (L), medium (M) and high (H). As was proved in the proposition, these differences are always positive, showing that the  $z$  functions never cross (and they are also ranked,  $z_L > z_M > z_H$ ). Thus, we should not worry about positive convergence points for the baseline

case. The main considerations for the selection of an upper bound  $\bar{\Phi}$  are *computational*, which are highly non-trivial and the subject of the next section.

**Different parameterizations.** To complete the discussion, it is useful to consider what happens for  $\rho \neq 1$ . For simplicity, consider again risk-sensitive preferences. The middle and the right panel in Figure C.3 plot the differences in the  $z$  functions for  $\rho = 0.5$  and  $\rho = 2$  respectively.<sup>2</sup> The middle panel shows that  $z$  functions do not cross for  $\rho = 0.5$ , so the only consideration for the selection of the upper bound is again computational. The right panel considers the most interesting case of  $\rho > 1$ . As we showed before, there is no upper bound in this case (marginal utility goes too quickly to infinity when consumption approaches zero). However, the graph shows that the ranking of the constant- $\Phi$  suboptimal positions  $z$  changes when the excess burden of taxation is high. At high levels of  $\Phi$ , the position against the low shock can become smaller than the position against the high shock (the difference in the graph becomes negative). In other words, fixing  $\Phi$ , when government spending increases, consumption falls and marginal utility increases so much that debt in marginal utility units increases. This happens for levels of  $\Phi$  that correspond to high levels of debt (above 1000% debt-to-output ratio as shown in figure C.2). Another crucial point about the right panel is the following: the differences in  $z$  do not cross the x-axis at the same value of  $\Phi$ . This means that there is no candidate convergence point when shocks take more than two values.<sup>3</sup>

Consequently, the analysis for  $\rho = 2$  suggests two things: first, the upper bound  $\bar{\Phi}$  should be large enough in order to include the points where the suboptimal  $z$  functions cross. Second, for such an upper bound, we can conjecture the following: if the *optimal* policy functions for  $z$  behave – with respect to their ranking – in a similar way as the suboptimal ones plotted in the graph, then a stationary distribution *may* exist without having to resort to an ad-hoc bound: there will be a positive drift in the excess burden with respect to the physical measure for low levels of debt, and a negative drift in the excess burden for very high levels of debt in the long-run (when the ranking of the policy functions changes, the covariance term in proposition 3 in the text becomes negative). Overall, we do not expect the short- and medium-run results to be substantially different than what the baseline analysis suggests. The long-run results though may be very different, especially when sufficient debt is accumulated.

To conclude, the above analysis suggests an easy procedure that can be used also for different utility functions like the balanced growth preferences of Chari et al. (1994): derive the induced  $z$  positions from a suboptimal constant- $\Phi$  policy and construct a state space that is large enough to

---

<sup>2</sup>In the EZW case, the object that determines the variation in the relative debt position  $\eta_t$  is the continuation- and marginal-utility adjusted debt position. So we would consider plots of  $V_t^{\rho-1} z_t$ , that are induced by a constant- $\Phi$  policy.

<sup>3</sup>If the shock took two values, then the point where the difference becomes zero would be a fixed point of the law of motion for the excess burden of taxation, since it implies a zero relative debt position  $\eta$ . But this would just be an artifact of an arbitrary restriction on the number of realizations of shocks.

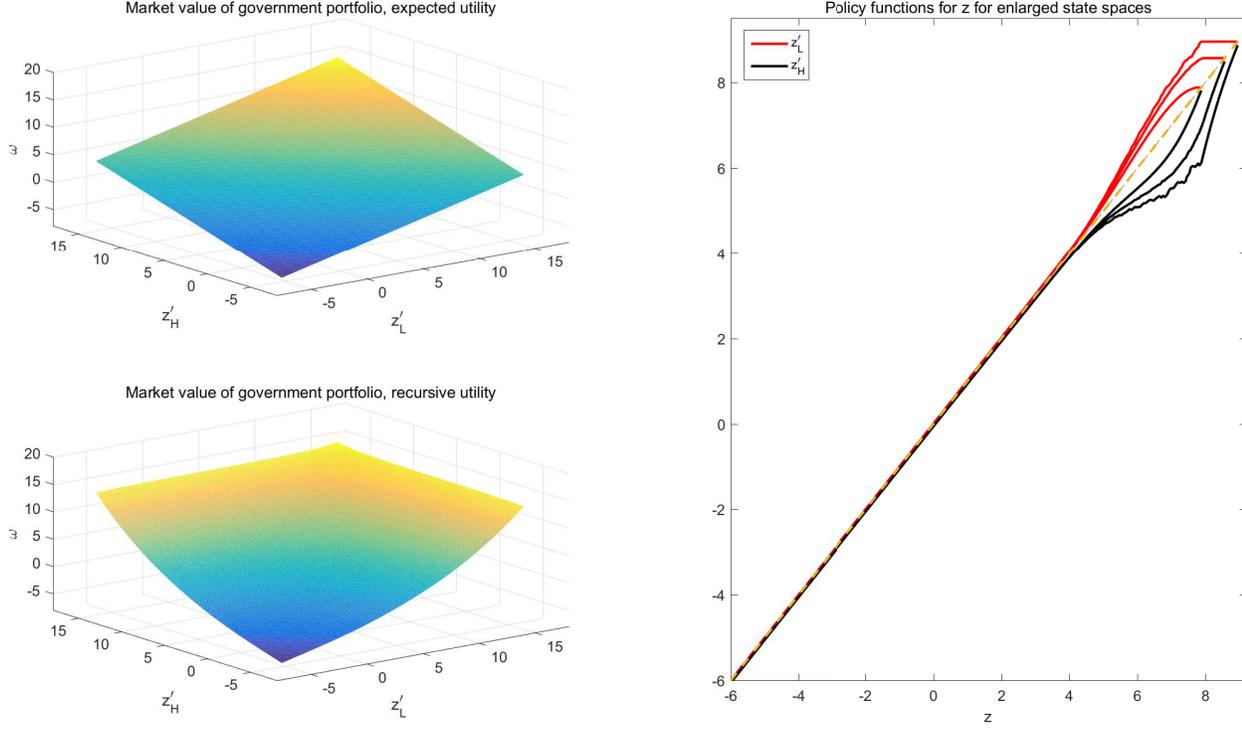


Figure C.4: The market value of the government portfolio is a convex function of the positions  $z'_i, i = L, H$  in the recursive utility case. The graph in the right plot the policy functions for  $z'_i$  for the baseline state space ( $\bar{\Phi} = 0.5$ ) and the enlarged ones.

Table C.1: Upper bounds of state space for the baseline exercise.

	$\bar{\Phi} = 0.5$	$\bar{\Phi} = 0.55$	$\bar{\Phi} = 0.58$
$\bar{\tau} \equiv \frac{\bar{\Phi}(\rho + \phi_h)}{1 + \bar{\Phi}(1 + \phi_h)}$ in %	50	52.38	53.70
$\bar{z}_L$ ( $b/y$ in %)	7.8819 (593.53)	8.5708 (640.93)	8.9546 (666.89)
$\bar{z}_H$ ( $b/y$ in %)	7.8391 (554.56)	8.5288 (598.27)	8.9129 (622.11)

The lower bounds of the state space are  $(\underline{z}_L, \underline{z}_H) = (-6.2355, -6.2911)$ , which correspond to *asset-to-output* ratios that are 510.06% and 492.23% respectively.

include points where these functions cross each other for different values of the shocks (if they do).

The entire discussion is based on the assumption that the computation of the Ramsey problem is a feasible task. The next section considers particular problems that may emerge.

### C.2.3 Non-convexities and robustness checks for the baseline exercise

**Non-convexities.** The upper bounds of the state space have to be chosen in a judicious way and may require some experimentation in order to make the computation of the problem feasible and the numerical results credible. The main difficulty is coming from novel non-convexities in the implementability constraint that may lead to jumps in policy functions or even to non-convergence issues. The non-convexities come from the surplus in marginal utility units,  $\Omega \equiv U_c c - U_l h$ , – a standard potential non-convexity in the time-additive setup –, and from the market value of the government portfolio  $\omega_t \equiv E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$ . For the particular utility function in the baseline exercise,  $\Omega$  is concave in  $(c, h)$ . Furthermore, the market value of the government portfolio is *linear* in  $z_i$  in the time-additive case. Thus, the particular Ramsey problem we solve is actually convex when  $\gamma = \rho = 1$ . However, with recursive utility,  $\rho = 1 < \gamma$ , even when we have a concave  $\Omega$ ,  $\omega$  becomes a *convex* function of  $z_i$  as figure C.4 shows (we need *concavity* of  $\omega$  to guarantee a convex constraint set). Non-convexities in  $\omega$  become *stronger* when a) the state space is increased b) the deviation from expected utility is larger, i.e. the difference of  $\gamma - \rho > 0$  and c) the size/volatility of shocks is increased, since all these factors increase the quantitative importance of continuation values in the determination of  $\omega$ .

I generated the state space by picking an upper bound that corresponds to  $\bar{\Phi} = 0.5$ . The respective tax rate is 50% and the maximum debt-to-output ratio is pretty large, of the order of 550 – 600%. In order to check the robustness of the results, I recalculated the problem for state spaces that correspond to  $\bar{\Phi} = 0.55, 0.58$ . Table C.1 reports the respective upper bounds. The right graph in figure C.4 shows the corresponding policy functions. What is interesting to observe is that when the state space becomes larger, the planner is taking larger positions against low shocks next period,  $z'_L$ , and smaller positions against large shocks,  $z'_H$ , even if he is at parts of the state space, for which he was not originally constrained. Thus, the fiscal hedging and the overinsurance of the planner are even stronger and the policy functions change in a non-trivial way. For large state spaces, the non-convexities become stronger and the policy functions start having small jumps (which become larger for larger state spaces), leading to convergence issues.

**Medium-run.** Table C.2 reports ensemble moments of 10000 sample paths of 2000 period length of the tax rate and the debt-to-output ratio for the baseline case of  $\bar{\Phi} = 0.5$  (that correspond to figure 3 in the text) and for the enlarged state spaces. The same realization of shocks was used across the three simulations. The ensemble moments for this sample length across the three different state spaces show small differences of 1-2 basis points for the mean and 5-10 basis points for the standard deviation of the tax rate. The mean debt-to-output ratio may differ up to 40 basis points and the standard deviation up to 90 basis points. We conclude that medium term statistics are robust to small increases in the state space.

Table C.2: Ensemble moments for larger state spaces.

	t=200	t=500	t=1000	t=1500	t=2000
$\bar{\Phi} = 0.5$ (baseline)					
<b>Tax rate in %</b>					
Mean	22.38	22.47	22.62	22.76	22.93
Standard deviation	0.43	0.69	0.99	1.25	1.48
95th percentile	23.12	23.68	24.34	24.96	25.54
5th percentile	21.66	21.4	21.08	20.88	20.73
<b>Debt-to-output ratio in %</b>					
Mean	-0.57	1.57	4.72	7.75	11.37
Standard deviation	9.68	15.37	21.83	27.44	32.40
95th percentile	15.17	27.51	42.32	55.98	68.39
5th percentile	-16.51	-22.21	-29.05	-33.45	-36.78
$\bar{\Phi} = 0.55$					
<b>Tax rate in %</b>					
Mean	22.37	22.46	22.60	22.74	22.91
Standard deviation	0.43	0.71	1.02	1.3	1.55
95th percentile	23.12	23.68	24.37	25.03	25.67
5th percentile	21.66	21.37	21.03	20.81	20.63
<b>Debt-to-output ratio in %</b>					
Mean	-0.78	1.19	4.24	7.27	10.96
Standard deviation	9.75	15.64	22.49	28.53	33.89
95th percentile	15.33	28.19	43.43	58.10	71.55
5th percentile	-16.74	-22.77	-30.12	-35.01	-38.78
$\bar{\Phi} = 0.58$					
<b>Tax rate in %</b>					
Mean	22.37	22.46	22.60	22.74	22.91
Standard deviation	0.43	0.70	1.01	1.28	1.52
95th percentile	23.12	23.68	24.35	25.00	25.63
5th percentile	21.66	21.39	21.05	20.83	20.67
<b>Debt-to-output ratio in %</b>					
Mean	-0.72	1.28	4.32	7.32	10.97
Standard deviation	9.70	15.49	22.18	28.07	33.29
95th percentile	15.12	27.83	42.83	57.34	70.52
5th percentile	-16.65	-22.54	-29.68	-34.37	-37.98

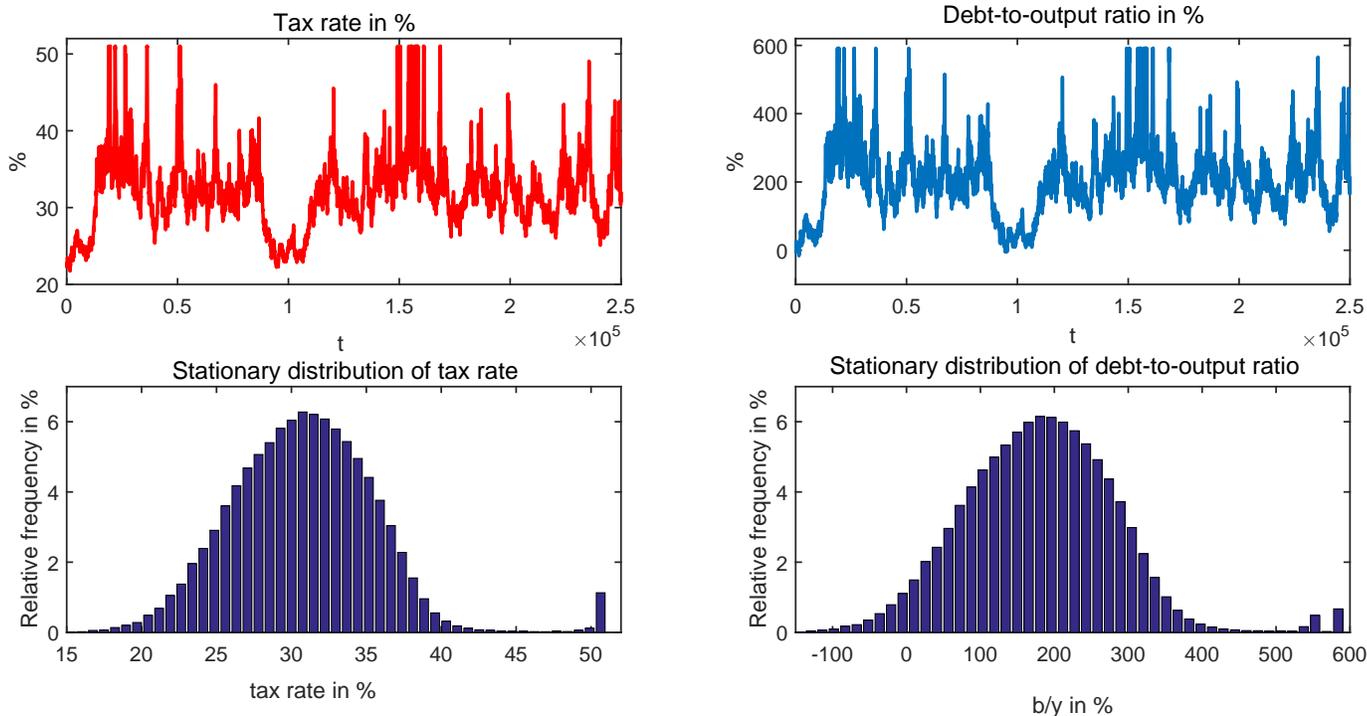


Figure C.5: The top panel displays the tax rate and the debt-to-output ratio for the first 250,000 periods of a 60 million long simulation. The lower panel displays the stationary distribution of these variables. The first 2 million observations were dropped.

**Long-run.** It remains to be seen if there is a lot of probability mass in the upper parts of the state space where the policy functions change in a substantial way when the state space is enlarged. Figure C.5 provides information for the long-run simulation. The upper panels show how the positive drifts of the tax rate and the debt-to-output ratio break down when the upper bound of the state space is hit and the lower panels plot the respective stationary distributions. Figure C.6 contrasts the stationary distributions of the tax rate and the debt-to-output ratio for the baseline and enlarged state spaces and table C.3 reports the respective moments. The change in the moments of the tax rate across state spaces is small, whereas for the debt-to-output ratio is more noticeable. Note that the larger the state space, the more *concentrated* the distribution is and the *thinner* the upper tail of the tax rate and the debt-to-output ratio. It is worth noting that even in the baseline state space, the 95th percentile of the tax rate is 37.93%. The 95th percentile of  $z$  is 4.42 and falls to 4 for the larger state space (recall from table C.1 that the upper bounds are 7.9 – 9). We conclude that the upper parts of the state space where policy functions change in a substantial way, are visited much less than 5% in the stationary distribution.

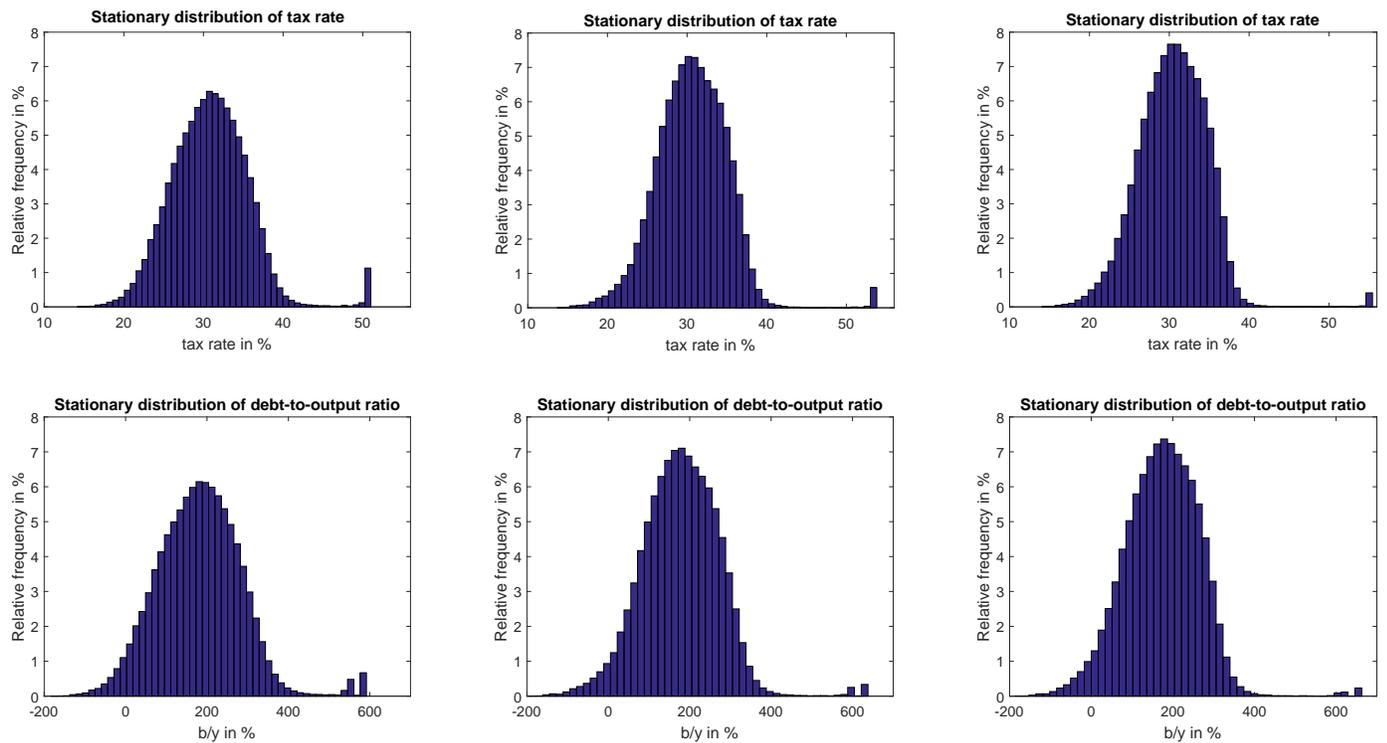


Figure C.6: Stationary distributions from a 60 million periods simulation for the three state spaces from left to right, ( $\bar{\Phi} = 0.5, 0.55, 0.58$ ). The first 2 million periods were dropped. The same realization of shocks was used across the different state spaces.

Table C.3: Moments of stationary distributions for different state spaces.

	$\bar{\Phi} = 0.5$	$\bar{\Phi} = 0.55$	$\bar{\Phi} = 0.58$
<b>Tax rate in %</b>			
Mean	30.8618	30.6680	30.4484
St. Dev.	4.9355	4.5061	4.3403
95th pct.	37.8903	36.9837	36.5080
98th pct.	40.5975	38.3723	37.6479
99th pct.	50.5929	40.1709	38.7354
<b>Debt-to-output ratio in %</b>			
Mean	181.9735	178.0870	173.4902
St. Dev.	104.2824	95.5117	92.1834
95th pct.	334.4611	315.0606	304.9756
98th pct.	397.3001	348.7182	333.2482
99th pct.	551.1968	393.8507	361.4129
<b>Debt in marginal utility units <math>z</math></b>			
Mean	2.3919	2.3376	2.2756
St. Dev.	1.3927	1.2720	1.2256
95th pct.	4.4096	4.1474	4.0110
98th pct.	5.2380	4.5785	4.3675
99th pct.	7.7684	5.1616	4.7275

Moments from the stationary distribution from a 60 million period simulation for the three state spaces from left to right, ( $\bar{\Phi} = 0.55, 0.55, 0.58$ ). The first 2 million periods were dropped. The same realization of shocks was used across the different state spaces.

Table D.1: Higher risk aversion or higher shock variance.

	<i>Expected utility</i>	<i>Recursive utility</i>				
		t=200	t=500	t=1000	t=1500	t=2000
<b>Higher risk aversion (<math>\gamma = 11</math>)</b>						
<b>Tax rate in %</b>						
Mean	22.30	22.39	22.49	22.65	22.80	22.99
Standard deviation	0	0.50	0.80	1.15	1.46	1.74
95th percentile		23.25	23.90	24.66	25.39	26.1
5th percentile		21.57	21.26	20.89	20.65	20.48
<b>Debt-to-output ratio in %</b>						
Mean	-1.907	-0.44	1.92	5.36	8.69	12.74
Standard deviation	1.907	11.12	17.76	25.31	32.02	38.13
95th percentile		17.80	31.82	49.82	66.02	80.61
5th percentile		-18.48	-25.12	-33.14	-38.32	-42.22
<b>Higher shock variance</b>						
<b>Tax rate in %</b>						
Mean	22.29	22.46	22.67	23.00	23.35	23.76
Standard deviation	0	0.65	1.07	1.58	2.07	2.55
95th percentile		23.61	24.56	25.80	27.09	28.41
5th percentile		21.40	21.05	20.67	20.44	20.29
<b>Debt-to-output ratio in %</b>						
Mean	-2.83	0.05	4.65	11.92	19.44	28.26
Standard deviation	2.83	14.69	23.73	34.75	45.47	55.81
95th percentile		24.33	45.58	73.14	101.54	129.68
5th percentile		-23.57	-30.89	-39.31	-44.17	-47.51

Ensemble moments for the case of higher risk aversion ( $\gamma = 11$ ) or the case of shocks with a standard deviation that corresponds to 3% of average first-best output. In order to avoid sample uncertainty, I use the same realizations of shocks as in table C.2.

## D Higher risk aversion or more volatile shocks

It is natural to conjecture that larger risk aversion (fixing  $\rho$  to unity) or more volatile fiscal shocks that need to be insured against, will lead to more pronounced differences from the expected utility plan. Table D.1 reports the ensemble moments for two experiments of interest. At the upper part of the table risk aversion is increased to  $\gamma = 11$ , keeping the rest of the calibration the same. At the lower part, the standard deviation of the shocks is increased, keeping the mean value of the shocks and the rest of the parameters the same. In particular, I set  $g_L = 0.068$  and  $g_H = 0.092$

which correspond now to 17% and 23% of average first-best output, so the standard deviation of the share of government spending in average first-best output is 3% (instead of the baseline 2%).

In both cases the increase over time of the ensemble moments of the tax rate and the debt ratio are larger than for the baseline case. For the higher risk aversion case the changes are small (since we only increased it to  $\gamma = 11$ ) but still noticeable. For the higher volatility case (which still remains comparable to the typical calibration for post war U.S. data), the changes are more noticeable. For example at  $t = 2000$  the tax rate has a standard deviation of 2.5%.

## E Persistent shocks and fiscal insurance exercise

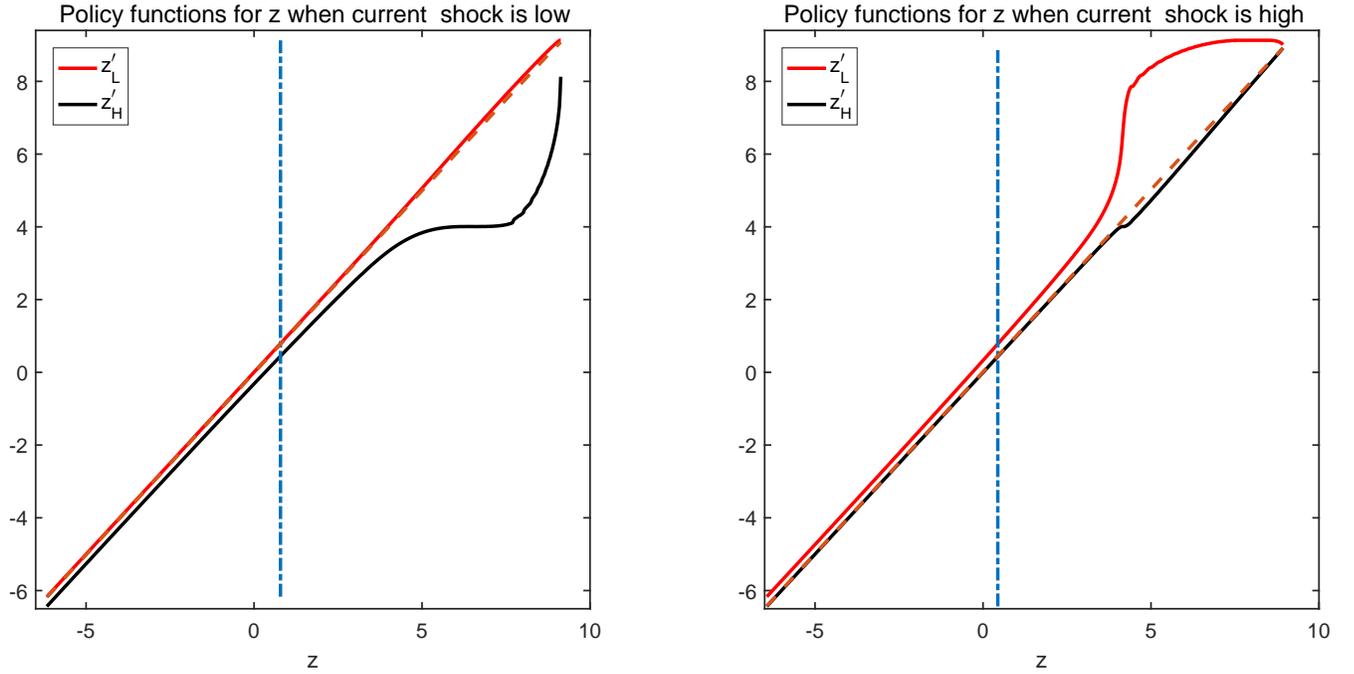


Figure E.1: Policy functions for debt next period, given a low or high current shock. The dotted line represents the market value of the government portfolio  $\omega$ . The vertical lines denote the position of the *optimal* initial value of the state,  $z_1$ .

### E.1 Calibration of shocks, policy functions and some sample paths

Consider now the fiscal insurance exercise. I set  $b_0 = 0.5 \times 0.4 = 0.2$  that corresponds to 50% of average first-best output. Let  $g_t = G \cdot \exp(x_t)$  where  $x_t = \rho_g x_{t-1} + \epsilon_t^g$  a zero-mean AR(1) process. As Benigno and Woodford (2006) and Farhi (2010), I follow the standard calibration of Chari et al. (1994),  $\rho_g = 0.89, \sigma_g = 0.07$ , where  $\sigma_g$  the unconditional volatility, which implies a conditional volatility of 0.0319. This captures well the dynamics of government consumption in post-war U.S. data.<sup>4</sup> I approximate the process with a symmetric Markov chain  $\pi_{ii} = 0.945, i = L, H$  and  $(x_L, x_H) = (-0.07, 0.07)$ . I set  $G = 0.2 \times \text{average first-best labor} = 0.08$ . This leads to  $(g_L, g_H) = (0.0746, 0.0858)$ . The choice of  $G$  leads to a share of government expenditures in first-best output with mean 20.04% and standard deviation 1.2451%. The respective share of government expenditures in output at the second-best expected utility economy has mean 22.71%

<sup>4</sup>Chari et al. (1994) use the annualized versions from Christiano and Eichenbaum (1992), who were using quarterly data for the period 1955:III-1983:IV. I use quarterly data for the period 1950:I- 2005:IV for nominal government consumption (NIPA Table 3.9.5 line 2) and deflate it with the respective price index (Table 3.9.4 line 2) in order to get real government consumption. The autocorrelation in linearly-detrended data is 0.9694 and the standard deviation 7.0157%, which leads to an annualized persistence parameter  $0.8831 \simeq 0.89$ , so I stick to the numbers of Chari et al. (1994).

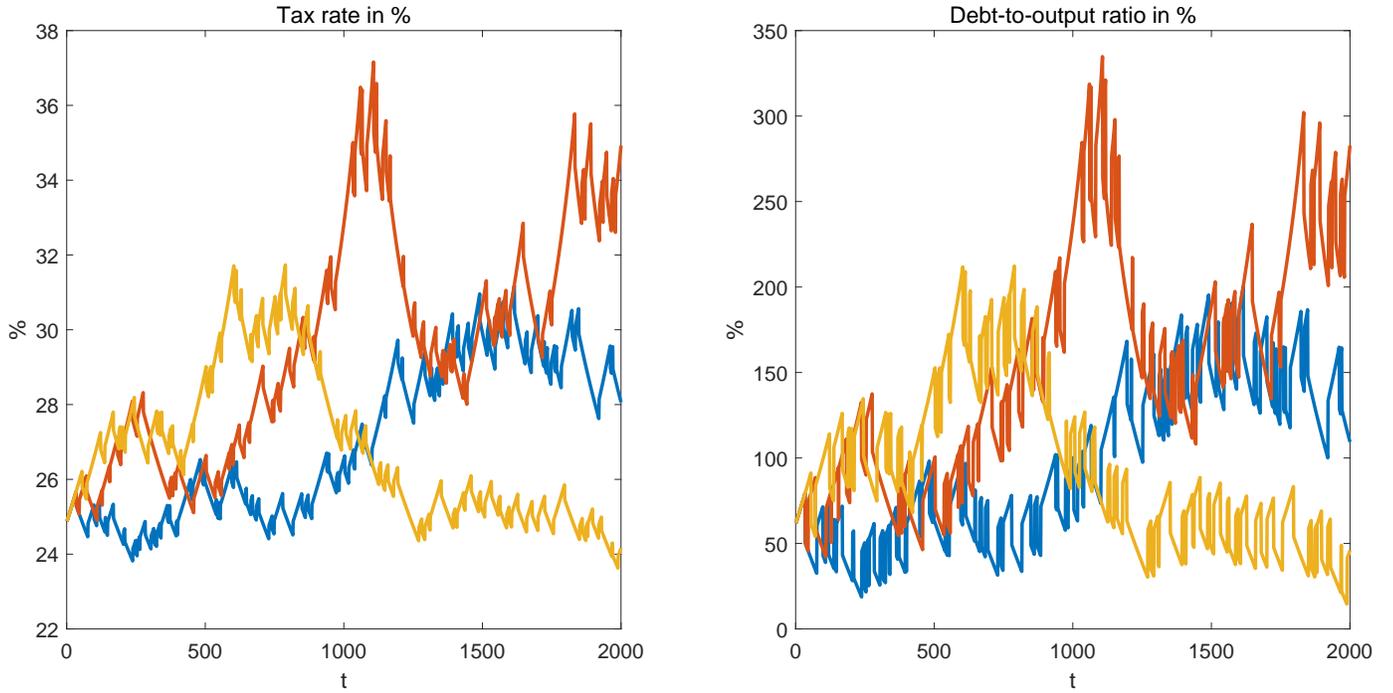


Figure E.2: Random sample paths of the tax rate and the corresponding debt-to-output ratio for the Chari et al. (1994) shock specification.

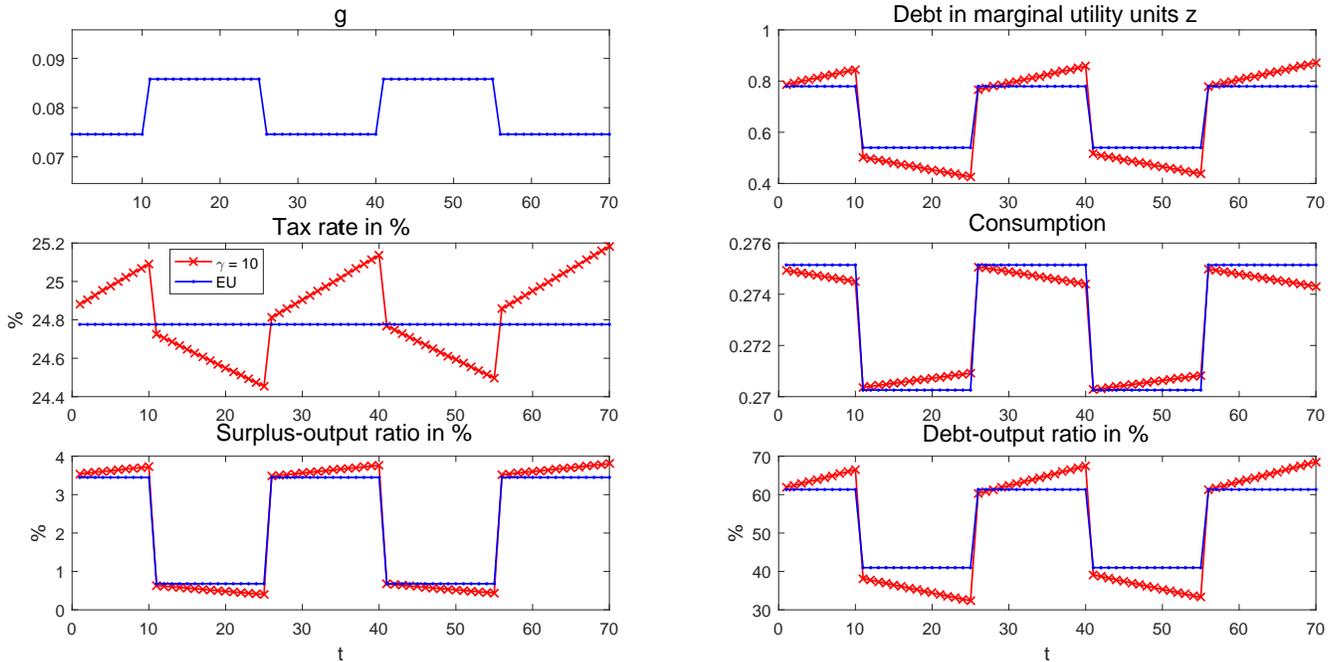


Figure E.3: Sample paths for an alternating sequence of low and high shocks for the Chari et al. (1994) government spending specification.

and standard deviation 1.38%.<sup>5</sup>

<sup>5</sup>Working with an economy with capital, Chari et al. (1994) were calibrating this parameter in order to get a share of government spending that is 16.7% in the deterministic steady state. For the period 1950:I- 2005:IV the

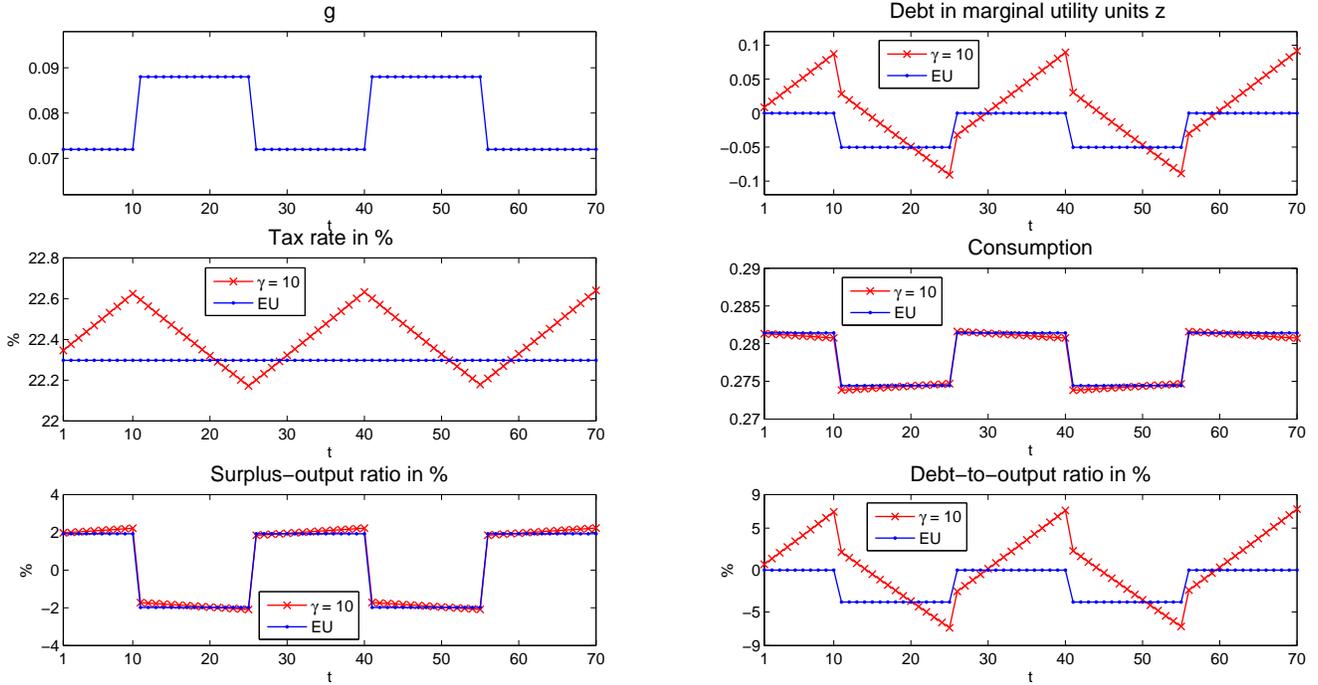


Figure E.4: Sample paths for an alternating sequence of low and high shocks for the baseline exercise.

The rest of the calibration is the same except for a small change in the labor disutility parameter, which is set to  $a_h = 7.8173$  so that the household works 40% at the first best. For the state space I set  $\bar{\Phi} = .59$  and use 800 points for each state space  $Z_i, i = L, H$ . For the final step of precision I fit cubic splines with 267 breakpoints.

Figure E.1 displays the policy functions for the persistent case. The graph shows that the relative debt position  $\eta_i = z'_i - \omega, i = L, H$  is becoming large only when there is a switch from a high to low shock (and the opposite). This is because the persistent shock is weighted heavily in the market value of debt,  $\omega$ , and therefore the relative debt position is large only when there is a transition to the other shock.

Figure E.2 plots sample paths of the tax rate and the corresponding debt ratio, in order to get a feel of the persistence and volatility of these variables. In order to understand the dynamics of the optimal plan, consider at  $t = 1$  a sample path of 10 low shocks, followed by a sequence of shocks that alternates between 15 high and 15 low shocks. Figure E.3 contrasts the expected and recursive utility plan for this shock realization. The planner is hedging fiscal shocks every period by taking a large state-contingent position against low shocks and a small against large shocks. Consequently, at each period that the shock remains low, the change in the tax rate is positive and the tax

---

share of government consumption in GDP is 16% with standard deviation of 1.14%. If we enlarge the notion of government purchases to consumption consumption expenditures plus gross investment we get a mean share in output of 21.2% with standard deviation of 1.95%. For the same period, the model-implied share, i.e. the share of government consumption to the model notion of output, i.e. government consumption/(non-durable goods plus services plus government consumption), has mean 23% and standard deviation 1.90%.

rate is increasing over time till the first switch. The debt position in marginal utility units is also increasing over time till the first switch, which translates to an increasing debt-to-output ratio.<sup>6</sup> When the shock switches to the high value the opposite pattern emerges. The government, which allocates less distortions on high shocks, starts reducing the tax rate over time. Debt in marginal utility units drops when the shock becomes high and then starts to decrease slowly reflecting the decrease of the tax rate. The opposite pattern emerges again when we switch to the low shock. Remember that in the expected utility case the tax rate would stay constant and that debt in marginal utility units would fluctuate across two values.<sup>7</sup> A similar picture obviously emerges for the baseline case, as figure E.4 shows. When shocks are i.i.d. though, changes in the tax rate are generally small. We will see that also in the next section where we contrast moments.

## E.2 Moments of interest

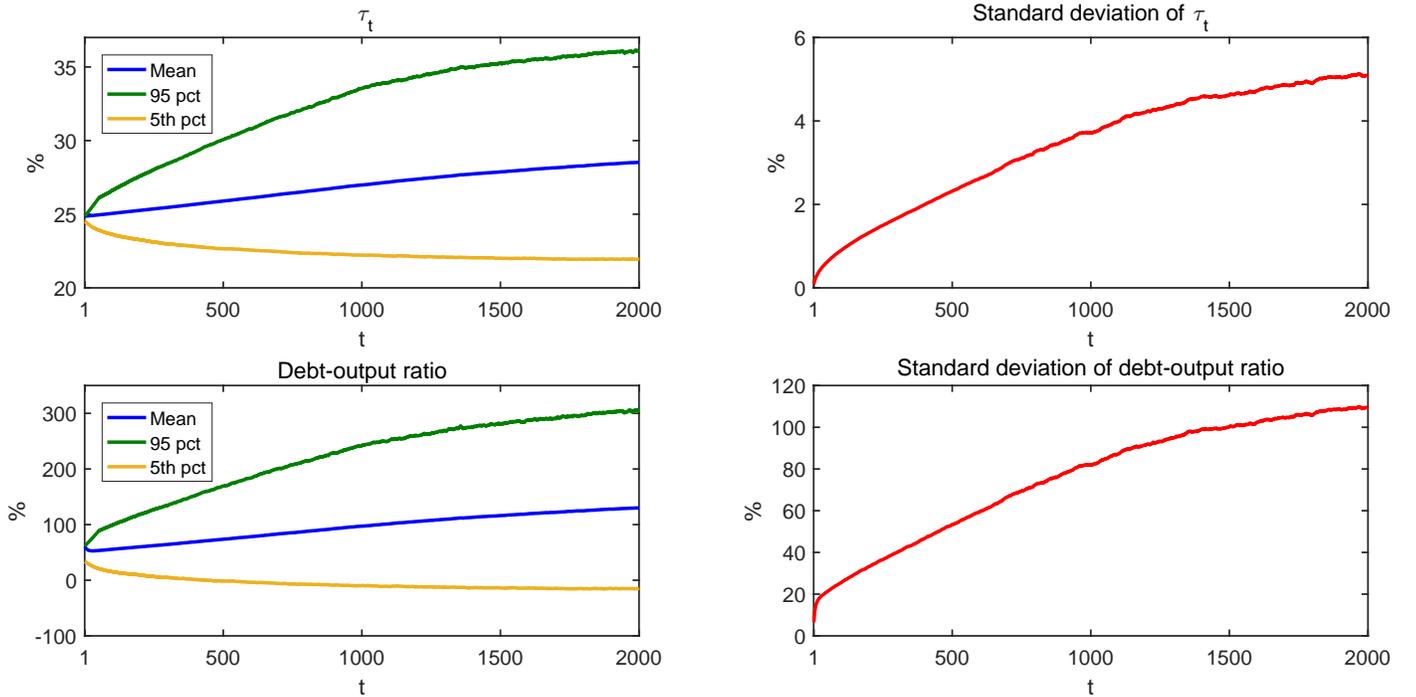


Figure E.5: Ensemble moments of the simulation with persistent shocks from 30,000 sample paths.

Figure E.5 displays the evolution of moments over time and table E.1 reports the exact numbers. What is important to note is that the positive drift and the standard deviation of the labor tax and the debt-to-output ratio are much *stronger* with persistent shocks and the stationary distribution

<sup>6</sup>The increase in the debt position in marginal utility units over time is an outcome of the numerical finding that the value functions are concave in  $z$  for each shock, and therefore the absolute value of the slope,  $\Phi_t$ , is increasing in  $z$ .

<sup>7</sup>Even if I used a period utility function that would imply a fluctuating tax rate in the expected utility case (for a example a utility function with time-varying Frisch elasticity), the tax rate would not change over time unless there was a switch in the shocks. This is due to the history-independence property.

Table E.1: Ensemble moments (persistent shocks).

	<i>Expected utility</i>	<i>Recursive utility</i>				
		t=200	t=500	t=1000	t=1500	t=2000
<b>Tax rate in %</b>						
Mean	24.78	25.26	25.89	26.98	27.87	28.52
St. Dev	0	1.32	2.32	3.72	4.62	5.09
95th pct		27.61	30.07	33.56	35.27	36.13
5th pct		23.26	22.65	22.23	22	21.93
<b>Debt-to-output ratio in %</b>						
Mean	51.15	60.02	73.69	97.25	116.21	130.14
St. Dev.	10.17	33.27	53.27	81.92	100.25	109.50
95th pct		118.64	169.33	242.39	281.61	307.51
5th pct		9	-1.40	-9.42	-13.92	-14.81

The table reports the expected utility moments and the ensemble moments with recursive utility that correspond to figure E.5.

Table E.2: Moments from the stationary distribution (persistent shocks).

<b>Stationary distribution</b>					
	$\tau$ in %		b/y in %		
Mean		30.49		172.15	
St. deviation		5.52		117.05	
95th pct		38.10		356.86	
98th pct		46.94		528.45	
St. deviation of change		0.4141		12.48	
Autocorrelation		0.9972		0.9943	
<b>Correlations (<math>\tau, b, g</math>)</b>					
Corr( $\Delta\tau, g$ )	-0.2649	Corr( $\Delta b, g$ )	-0.3069	Corr( $\Delta\tau, b$ )	0.0547
Corr( $\Delta\tau, \Delta g$ )	-0.5783	Corr( $\Delta b, \Delta g$ )	-0.8941	Corr( $\Delta\tau, \Delta b$ )	0.8316
Corr( $\tau, g$ )	-0.1804	Corr( $b, g$ )	-0.2737	Corr( $\tau, b$ )	0.9790

The simulation is 10 million periods long. The first million periods was dropped. The maximum tax rate is 57.4% and the maximum debt-to-output ratio 664.06%.

is reached *quicker*. The mean labor tax is 24.86% at  $t=1$  and becomes 28.5% at  $t = 2000$ , and the standard deviation from almost 0.08% becomes 5% at  $t = 2000$ . Similarly the mean and the standard deviation of the debt-to-output ratio become 130% and 110% respectively at  $t = 2000$ .

Table E.3: Comparison of moments for different shock specifications.

Tax rate in %	i.i.d.	CCK shocks	$2 \times \text{std}(g)$
Mean	30.86	30.49	31.26
St. Dev	4.94	5.52	7.76
St. Dev of $\Delta$	0.17	0.41	0.90
Autocorrelation	0.9994	0.9972	0.9932

The table contrasts the moments of the stationary distribution for the baseline calibration in the text (first column), the [Chari et al. \(1994\)](#) specification (second column) and the case where the unconditional volatility of the [Chari et al. \(1994\)](#) shocks is doubled (third column).

Table [E.2](#) reports the moments of the stationary distribution. The correlation  $\text{Corr}(\Delta\tau, g)$  is smaller than in the baseline case in the text, due again to the small changes in tax rates when the shock remains the same. However, the correlation of changes in tax rates and changes in state-contingent debt remains highly positive,  $\text{Corr}(\Delta\tau, \Delta b) = 0.83$ .

Finally, table [E.3](#) contrasts the moments from the baseline i.i.d. exercise to the moments coming from the persistent specification of [Chari et al. \(1994\)](#). It considers also the moments from an exercise where the unconditional volatility of the persistent specification is doubled (in that case government expenditures are larger but they still do not resemble war-peace shocks). What is interesting to observe is that by having persistent shocks (with a smaller but similar unconditional volatility as in the baseline case), the unconditional volatility of the tax rate may increase a little, but the standard deviation of the *change* in the tax rate,  $\Delta\tau$ , is more than doubled, reaching 41 basis points. So there is much more action with persistent shocks (which is why the stationary distribution is reached quicker). If we double also the volatility of the shocks, the standard deviation of the change in tax rates reaches 90 basis points, and the unconditional volatility 7.5 percentage points. Consequently, if we increase the risk in the economy (by increasing the persistence and volatility of shocks), the quantitative effects of recursive utility on the optimal taxation problem become even more pronounced.

### E.3 Returns and risk premia

The return on the government portfolio can be written as

$$R_{t+1} = \frac{b_{t+1}}{\sum_{g_{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1})} = \frac{b'_{g'}}{\beta\omega(z, g)/U_c(z, g)} = \frac{U_c(z, g)}{\beta\omega(z, g)} \cdot \frac{z'_{g'}(z, g)}{U_c(z'_{g'}(z, g), g')} = R(g', z, g).$$

Figure [E.6](#) plots the returns for the expected utility case ( $\gamma = 1$ ) versus the recursive utility case ( $\gamma = 10$ ). It shows that the government is reducing the return on the government debt when bad

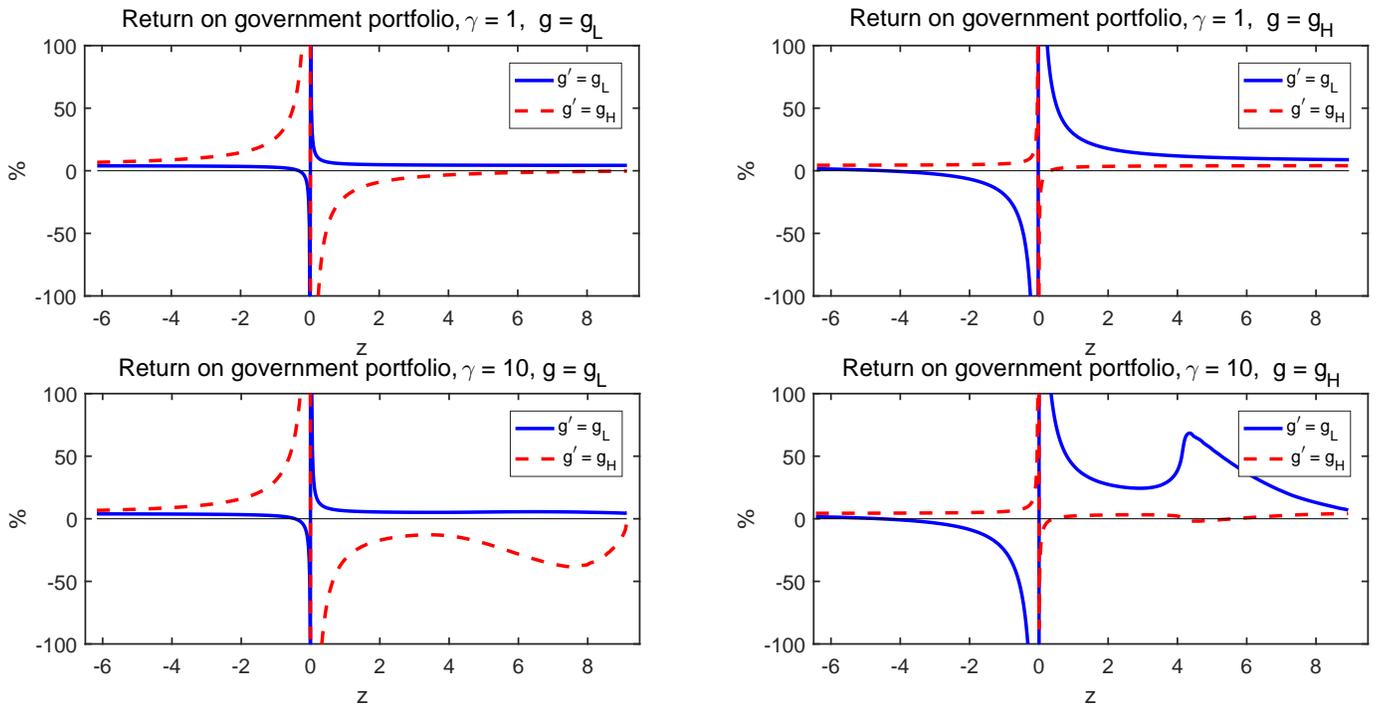


Figure E.6: The top panel displays the return on the portfolio of government debt for the time-additive case and the bottom panel for recursive utility. Government expenditure shocks are calibrated as in [Chari et al. \(1994\)](#).

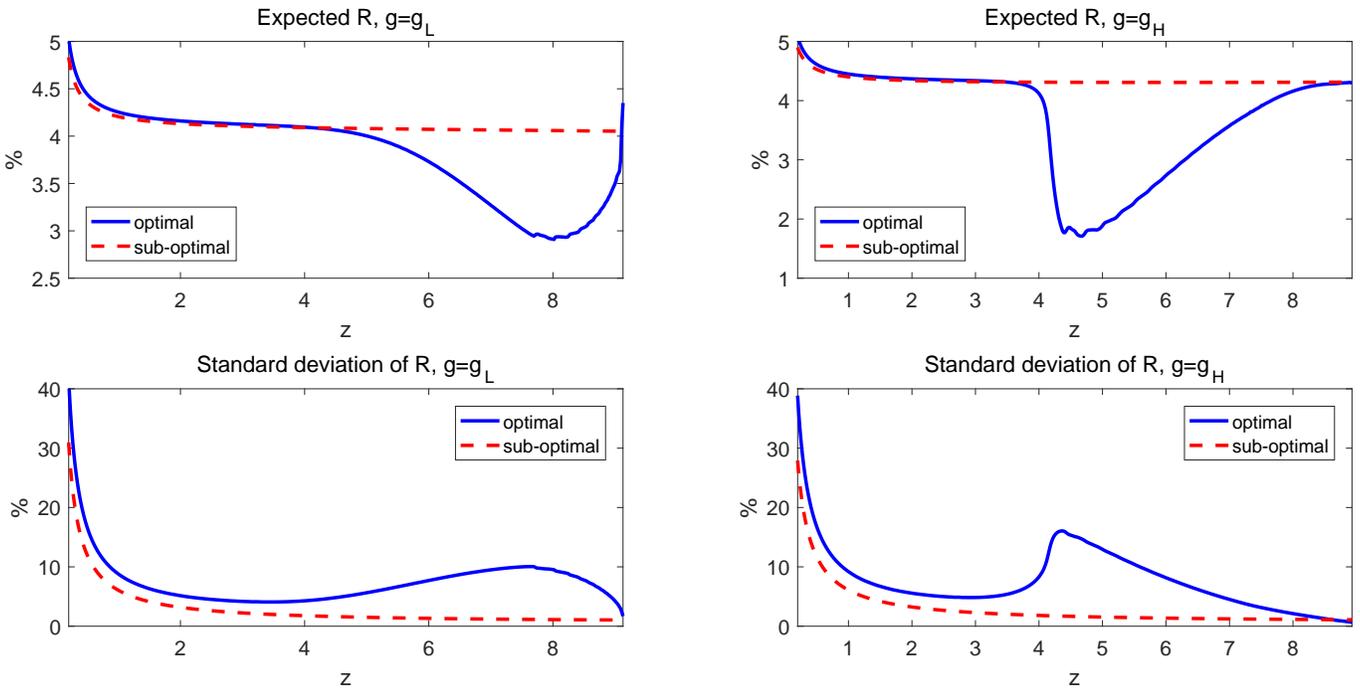


Figure E.7: Conditional mean and standard deviation of government portfolio returns resulting from the optimal policy and the constant  $\Phi$  policy.

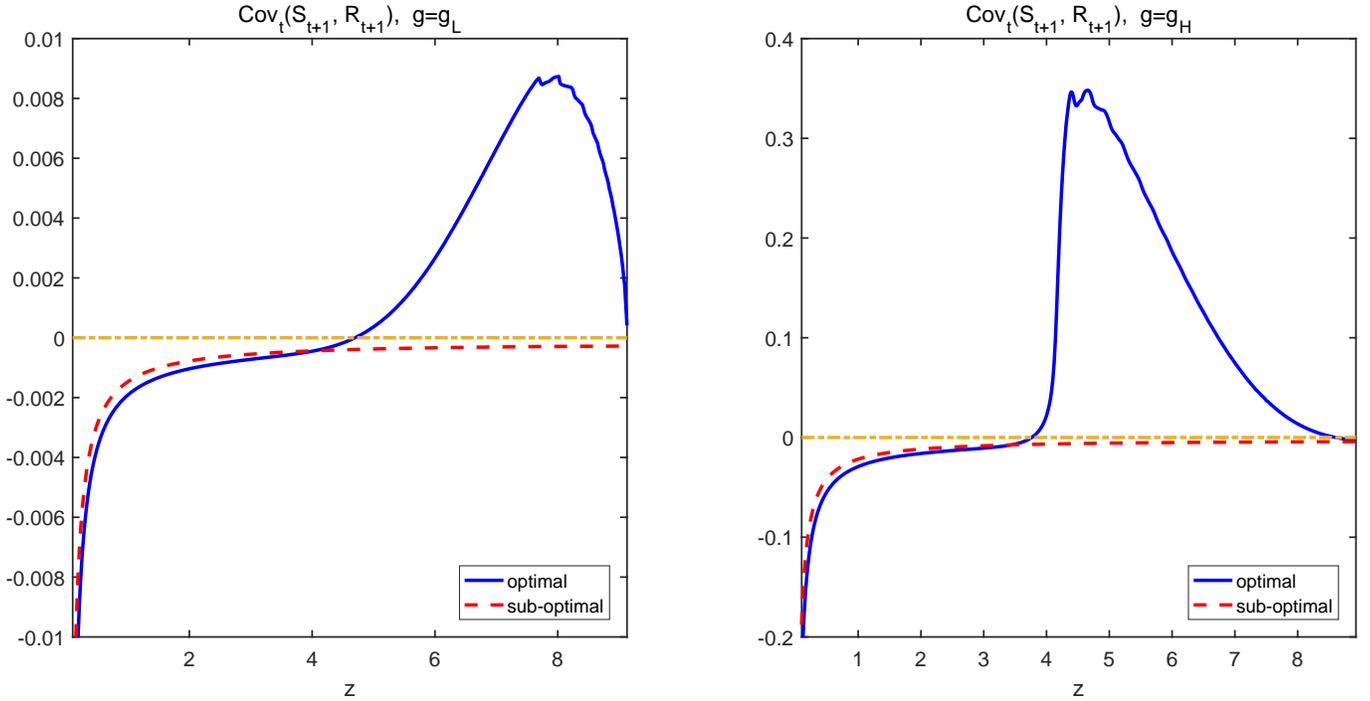


Figure E.8: Conditional covariance of the stochastic discount factor with the returns of the government debt portfolio.

shocks realize and compensates bond-holders with high returns when government expenditures become small again. The opposite happens when the government holds liabilities against the private sector ( $b_t < 0$ ), i.e. it increases the returns on assets in bad times and reduces them in good times (which can be seen at the left quadrants of each graph in figure E.6).

Figure E.7 contrasts the conditional mean and the standard deviation of the optimal debt returns to the suboptimal returns induced by constant excess burden policies. It shows how expected returns fall for large amounts of debt, which leads to the negative risk premium. Figure E.8 makes the same point by displaying the conditional covariance of the stochastic discount factor with debt returns. It starts negative and it becomes positive, making the government portfolio a hedge, when debt is high.

Table E.4 reports moments of the optimal returns, the risk free rate and the market price of risk in the two economies at the stationary distribution. In the calculation of returns in the recursive utility economy, I excluded abnormal returns by trimming realizations above the 99.5 percentile and below the 0.5 percentile. This is because when debt becomes close to zero or when it switches to negative, returns become abnormally high (of the order of 1000%) or abnormally negative, due to divisions with numbers that are close to zero, as can be seen at the vertical asymptotes of figure E.6. The last column in the table excludes situations with assets (i.e. liabilities of the private sector towards the government), because the BLY methodology accommodates only debt. The mean and standard deviation of optimal returns do not differ much across the two economies.

Table E.4: Returns of government portfolio and market price of risk.

	Expected utility	Recursive utility	
		no outliers	no outliers/no assets
<b>Return in %</b>			
Mean	4.2038	4.39	4.3338
St. Dev.	9.4426	8.4874	8.0105
<b>Risk-free rate in %</b>			
Mean	4.1658	4.1582	4.1569
St. Dev.	0.10266	0.1122	0.11262
<b>Excess return in %</b>			
Mean	0.037972	0.2318	0.1770
St. Dev.	9.4419	8.4875	8.0135
<b>SDF</b>			
Mean	0.96001	0.9601	0.9601
St. Dev.	0.0040344	0.0208	0.021
MPR	0.0042025	0.0216	0.0218
<b>Decomposition of variance of log SDF in %</b>			
St. Dev. of $\log(\text{SDF})$	0.42021	2.1584	2.1758
St. Dev. of $\Delta \log(c_{t+1})$	0.42021	0.4159	0.4209
St. Dev. of log EZ term	-	1.7709	1.7837

The simulation is 10 million periods long and the first million of observations was dropped. Outlier returns have been excluded from the calculation of the statistics for the recursive utility economy. Outliers are observations above the 99.5 percentile and below the 0.5 percentile. The probability of negative debt (i.e. assets) is 4.4% at the stationary distribution. The excluded observations with either outliers or assets are 4.91 % of the sample.

The big differences emerge in the market price of risk, which becomes five times larger, from 0.004 to 0.021, despite the fact that there is limited risk due to small fiscal shocks and the absence of any other risks like technology shocks. The increase in the market price of risk comes mainly from an increase in the standard deviation of the recursive utility term. The standard deviation of consumption growth is about 0.42% in both economies but the recursive utility term has a standard deviation of 1.77%, leading to an overall standard deviation of the logarithm of the stochastic discount factor of 2.15%.

Table E.5: Fiscal insurance in post-war U.S. data from [Berndt et al. \(2012\)](#).

	<i>Valuation channel</i>	<i>Non-defense surplus channel</i>
<b>Beta</b>	-0.3663	2.7962
Current	-0.0690	
Future	-0.2973	
<b>Fraction in %</b>	9.61	73.34
Current	1.81	
Future	7.8	

Fiscal adjustment betas and fiscal insurance fractions in [Berndt et al. \(2012\)](#) (Table 3, page 85). They do not decompose the surplus channel to current and future news. The empirical correlation of news to defense spending with news to debt returns and non-defense surpluses is  $-0.72$  and  $0.40$  respectively.

## E.4 Fiscal insurance in [Berndt et al. \(2012\)](#) and log-linear approximation constants

Table [E.5](#) reproduces the fiscal adjustment betas and fractions of [Berndt et al. \(2012\)](#) for the reader's convenience.

The log-linear approximation of [Berndt et al. \(2012\)](#) is based on the assumption of positive debt and on the assumption that the government is running on average a surplus. The approximation is around the unconditional means of the logarithmic tax-revenue-to-debt and government-expenditures-to-debt ratios. Define  $\mu_{\tau b} \equiv \exp(E(\ln T_t/b_t))$  and  $\mu_{gb} \equiv \exp(E(\ln g_t/b_t))$ . The approximation constants are

$$\begin{aligned}\mu_T &\equiv \frac{\mu_{\tau b}}{\mu_{\tau b} - \mu_{gb}} \\ \mu_g &\equiv \mu_T - 1 = \frac{\mu_{gb}}{\mu_{\tau b} - \mu_{gb}} \\ \rho_{\text{BLY}} &\equiv 1 + \mu_{gb} - \mu_{\tau b}\end{aligned}$$

It is assumed that  $\mu_{\tau b} > \mu_{gb}$  and that  $\mu_{\tau b} - \mu_{gb} < 1$ . Thus, the government is running on average a surplus, but this surplus is not large enough to pay back the entire stock of current liabilities. These two assumptions imply that  $\rho_{\text{BLY}} \in (0, 1)$ . The constant  $\rho_{\text{BLY}}$  stands effectively for the value of the newly issued debt as a fraction of current liabilities (and is smaller than unity since the government is running a primary surplus instead of a deficit). The parameters  $(\mu_T, \mu_g)$  can be interpreted as average tax revenues and government expenditures as fractions of the surplus. As such, they both decrease when tax revenues are large, which is the case in an economy with recursive utility.

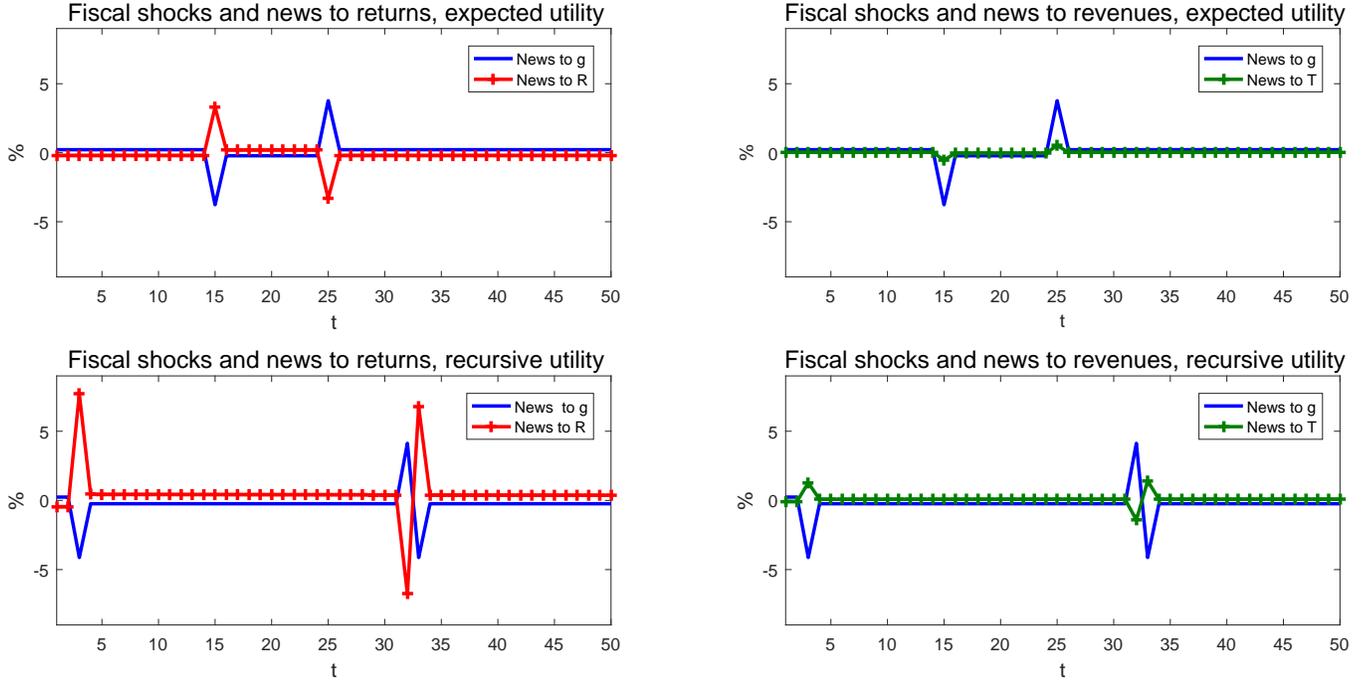


Figure E.9: BLY decomposition of fiscal shocks to news in returns and news to revenues for expected and recursive utility.

## E.5 News variables

**Expected utility.** In order to calculate the news/surprise variables in the decomposition of the intertemporal budget constraint, we are going to use properties of the history independence of the [Lucas and Stokey \(1983\)](#) allocation. Assume that we need to calculate the variable

$$I_{t+1} \equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho^i \Delta y_{t+i+1} = \sum_{i=0}^{\infty} \rho^i (E_{t+1} - E_t) \Delta y_{t+i+1}$$

where  $y_t = y(g_t)$  and  $g_t$  Markov with transition matrix  $\Pi$ . Let  $e_{g_t}$  denote the column vector with unity at the position of the shock at  $t$ ,  $g_t$ , and zero at the rest of the rows. Collect the values of  $y$  in the column vector  $\mathbf{y}$ . We have  $y_t = e'_{g_t} \mathbf{y}$  and  $E_t y_{t+i} = e'_{g_t} \Pi^i \mathbf{y}$ . After some algebra and using properties of discounted sums of Markov matrices, we get

$$I(g_{t+1}|g_t) = (e'_{g_{t+1}} I - e'_{g_t} \Pi) [I + \rho(I - \rho \Pi)^{-1} (\Pi - I)] \cdot \mathbf{y} \quad (\text{E.1})$$

This formula can be used for the calculation of news in the growth rate of fiscal shocks and news in the growth rate of tax revenues.

Consider now the case of returns, that are described by a matrix  $\mathbf{R} \equiv [R(g'|g)]$ , where  $R$

denotes now the logarithmic return. Define  $x_t \equiv E_t \sum_{i=0}^{\infty} \rho^i R_{t+1+i}$ , which satisfies the recursion

$$x_t = E_t R_{t+1} + \rho E_t x_{t+1}.$$

We have  $x_t = x(g_t)$ , and

$$x(g) = \sum_{g'} \pi(g'|g) R(g'|g) + \rho \sum_{g'} \pi(g'|g) x(g'), \forall g.$$

This recursion delivers the system

$$\begin{aligned} \mathbf{x} &= \begin{pmatrix} \sum_g \pi(g|1) R(g|1) \\ \dots \\ \sum_g \pi(g|N) R(g|N) \end{pmatrix} + \rho \Pi \mathbf{x} \\ &= (\Pi \circ \mathbf{R}) \cdot \mathbf{1} + \rho \Pi \mathbf{x}, \end{aligned}$$

where  $\mathbf{1}$  the  $N \times 1$  unit vector. Thus,

$$\mathbf{x} = (I - \rho \Pi)^{-1} (\Pi \circ \mathbf{R}) \cdot \mathbf{1}.$$

Define now  $y_{t+1} \equiv E_{t+1} \sum_{i=0}^{\infty} \rho^i R_{t+i+1}$  (this is a different  $y$  than the vector defined in (E.1)). We want to calculate the *surprise* in returns,  $I_{t+1}^R \equiv y_{t+1} - x_t$ . We have  $y_{t+1} = R_{t+1} + \rho E_{t+1} y_{t+2}$ , which implies

$$y(g'|g) = R(g'|g) + \rho \sum_{\hat{g}} \pi(\hat{g}|g') y(\hat{g}|g') \tag{E.2}$$

Let the matrix  $\mathbf{Y} \equiv [y(g'|g)]$  collect the unknowns  $y(g'|g)$ . From recursion (E.2) we get that each  $j$  column vector of  $\mathbf{Y}$  satisfies the system

$$\begin{pmatrix} y(j|1) \\ \dots \\ y(j|N) \end{pmatrix} = \begin{pmatrix} R(j|1) \\ \dots \\ R(j|N) \end{pmatrix} + \rho \mathbf{1} e_j' \Pi \mathbf{Y}' e_j, j = 1, \dots, N$$

Putting the  $N$  systems together delivers

$$\mathbf{Y} = \mathbf{R} + \rho[\mathbf{1}e'_1\Pi\mathbf{Y}'e_1, \dots, \mathbf{1}e'_N\Pi\mathbf{Y}'e_N]$$

In order to solve for  $\mathbf{Y}$ , we need to use the *vec* operator. After a lot of algebra we get

$$\text{vec}(\mathbf{Y}) = \text{vec}(\mathbf{R}) + \rho \cdot A \text{vec}(\mathbf{Y}) \Rightarrow \text{vec}(\mathbf{Y}) = [I_{N^2 \times N^2} - \rho A]^{-1} \text{vec}(\mathbf{R}),$$

where

$$A \equiv \begin{pmatrix} e'_1\Pi \otimes [\mathbf{1}_{N \times 1}, \mathbf{0}_{N \times 1}, \dots, \mathbf{0}_{N \times 1}] \\ e'_2\Pi \otimes [\mathbf{0}_{N \times 1}, \mathbf{1}_{N \times 1}, \dots, \mathbf{0}_{N \times 1}] \\ \dots \\ e'_N\Pi \otimes [\mathbf{0}_{N \times 1}, \mathbf{0}_{N \times 1}, \dots, \mathbf{1}_{N \times 1}] \end{pmatrix}$$

The matrix  $A$  is a partitioned matrix of dimension  $N^2 \times N^2$  with blocks of dimension  $N \times N^2$ .

**Recursive utility.** The previous formulas cannot be used for the recursive utility economy because of the dependence on the state  $z$ . To calculate the surprise in returns,  $I_{t+1} \equiv y_{t+1} - x_t$ , we need to solve numerically for the functions  $x$  and  $y$  from the following two recursions:

$$\begin{aligned} x(z, g) &= \sum_{g'} \pi(g'|g)R(g', z, g) + \rho \sum_{g'} \pi(g'|g)x(z'_{g'}(z, g), g') \\ y(g', z, g) &= R(g', z, g) + \rho \sum_i \pi(i|g')y(i, z'_{g'}(z, g), g') \end{aligned}$$

Similar calculations are necessary for the news in tax revenues. To conclude, figure E.9 plots sample paths of news to spending and news to returns and surpluses. News to tax revenues are positively correlated with news to spending for expected utility and *negatively* correlated for recursive utility.

## E.6 Linear approximation

In order to be able to use the log-linear methodology of [Berndt et al. \(2012\)](#), I ignored observations with negative debt that had probability 4.4% at the stationary distribution. In this section, I approximate the government budget constraint linearly, which allows the inclusion of  $b_t < 0$ . The right-hand-side of equation (28) in the text can be approximated as

Table E.6: News to expenditures, returns and revenues (linear).

	Expected utility			Recursive utility		
	$I^g$	$I^R$	$I^T$	$I^g$	$I^R$	$I^T$
$I^g$	0.0175			0.0169		
$I^R$	-0.9848	0.0158		-0.5673	0.0552	
$I^T$	1	-0.9848	0.0024	-0.6796	0.3602	0.0163

Standard deviations (on the diagonal– not multiplied with 100) and correlations of the news variables in the linear approximation.

Table E.7: Fiscal insurance (linear).

	Expected utility		Recursive utility	
	<i>Valuation channel</i>	<i>Surplus channel</i>	<i>Valuation channel</i>	<i>Surplus channel</i>
<b>Beta</b>	-0.8869	0.1398	-1.8482	-0.6557
Current	-0.9191	0.0204	-1.9145	-0.0106
Future	0.0322	0.1194	0.0663	-0.6451
<b>Fraction in %</b>	88.69	13.98	184.82	-65.57
Current	91.91	2.04	191.45	-1.06
Future	-3.22	11.94	-6.63	-64.51

Fiscal adjustment betas and fiscal insurance in the linear approximation. The approximation constants are  $(\bar{B}, \rho_{\text{linear}}) = (0.1730, 0.9597)$  and  $(\bar{B}, \rho_{\text{linear}}) = (0.5539, 0.9579)$  for the expected and recursive utility case respectively.

$$b_{t+1} = \text{const.} + \bar{B}R_{t+1} + \bar{R}[b_t + g_t - T_t]$$

where  $\bar{B} \equiv \bar{b} + \bar{g} - \bar{T}$ , and  $(\bar{R}, \bar{b}, \bar{T}, \bar{g}) = (E(R), E(b), E(T), E(g))$ , the respective unconditional means. Thus,  $\bar{B}$  stands for the average market value of new debt that the government has to issue in order to finance the primary deficit,  $g_t - T_t$ , and past obligations,  $b_t$ . Ignore the constant and rewrite the budget constraint as

$$b_t = T_t - g_t - \rho_{\text{linear}}\bar{B}R_{t+1} + \rho_{\text{linear}}b_{t+1},$$

where  $\rho_{\text{linear}} \equiv \bar{R}^{-1}$ . Solve forward, take expectations and use an asymptotic condition to get

$$b_t = E_t \sum_{i=0}^{\infty} \rho_{\text{linear}}^i (T_{t+i} - g_{t+i}) - E_t \sum_{i=1}^{\infty} \rho_{\text{linear}}^i \bar{B} R_{t+i}$$

Update one period, take expectation with respect to information at  $t$  and calculate the news (surprises) as

$$b_{t+1} - E_t b_{t+1} + (E_{t+1} - E_t) \sum_{i=1}^{\infty} \rho_{\text{linear}}^i \bar{B} R_{t+1+i} = (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{linear}}^i (T_{t+i+1} - g_{t+i+1})$$

Use the fact that a surprise in debt is a (scaled) surprise in returns,  $b_{t+1} - E_t b_{t+1} = \bar{B}(R_{t+1} - E_t R_{t+1})$ , to finally get

$$I_{t+1}^{g,\text{linear}} = -I_{t+1}^{R,\text{linear}} + I_{t+1}^{T,\text{linear}},$$

where

$$\begin{aligned} I_{t+1}^{g,\text{linear}} &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{linear}}^i g_{t+i+1} \\ I_{t+1}^{R,\text{linear}} &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{linear}}^i \bar{B} R_{t+i+1} \\ I_{t+1}^{T,\text{linear}} &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{linear}}^i T_{t+i+1}. \end{aligned}$$

Thus, news in the present value of spending can be decomposed as news in returns and news in tax revenues, furnishing the same interpretation as in the text, without restricting attention to the case of positive debt. We get  $1 = -\beta_R^{\text{linear}} + \beta_T^{\text{linear}}$ , where  $\beta_i^{\text{linear}}, i = R, T$  the respective fiscal adjustment betas. The valuation channel (in %) is  $-100 \cdot \beta_R^{\text{linear}}$  and the surplus channel is  $100 \cdot \beta_T^{\text{linear}}$ . Table E.7 compares fiscal insurance in the expected utility case and the recursive utility case using the linear approximation and delivers the same result as the BLY log-linear method: the valuation channel absorbs a much larger fraction of the shocks in the recursive utility economy (185%) and the surplus channel is negative (-65%). The size of the two channels is similar to the one reported in the text, so the exclusion of negative debt was not affecting the results in any substantial way. The only difference is in the decomposition of the surplus channel in terms of a current and future channel. In the linear decomposition, the future surplus channel (which calculates surprises in levels) is much more active than in the log-linear decomposition (which

Table E.8: Effect of approximation constants.

	Expected utility (log-linear)		Expected utility (linear)	
	<i>Valuation channel</i>	<i>Surplus channel</i>	<i>Valuation channel</i>	<i>Surplus channel</i>
<b>Beta</b>	-3.0074	0.6096	-0.8364	0.1671
Current	-3.3491	2.1988	-0.9286	0.0243
Future	0.3417	-1.5892	0.0922	0.1428
<b>Fraction in %</b>	83.49	16.92	83.64	16.71
Current	92.98	61.04	92.86	2.43
Future	-9.49	-44.12	-9.22	14.28

Fiscal insurance fractions in the expected utility case when initial debt is equal to the mean of the stationary distribution of debt with recursive preferences. We set  $b_0 = 0.6063$  for the log-linear exercise (where return outliers and assets are excluded) and  $b_0 = 0.5771$  in the linear exercise here (where only return outliers are excluded). The corresponding approximation constants are  $(\mu_g, \mu_\tau, \rho_{BLY}) = (3.6022, 4.6022, 0.9595)$  (log-linear case) and  $(\bar{B}, \rho_{\text{linear}}) = (0.5048, 0.9599)$  (linear case). These are similar to the approximation constants in the recursive utility case.

calculates surprises in growth rates).

## E.7 Effect of approximation constants

The level of average debt in the expected utility case is effectively determined by the level of initial debt. In contrast, in the recursive utility economy, long-run debt does not depend on the initial conditions and assumes pretty high levels. As a result, the approximation constants in either the log-linear  $(\mu_g, \mu_T, \rho_{BLY})$  or in the linear approximation  $(\bar{B}, \rho_{\text{linear}})$  are very different when we switch to the recursive utility case. This could lead to the concern that the expected utility numbers are somehow biased, due to low initial debt. Table E.8 performs the following thought experiment: it calculates the valuation and surplus channel for an expected utility economy with initial debt equal to the mean of the stationary distribution of the economy with recursive preferences (which started with low initial debt). This choice of initial debt leads to approximation constants in the expected utility economy that are similar to the approximation constants in the recursive utility economy. Still, the levels of the valuation and the surplus channel with expected utility do not change substantially. With high initial debt, the valuation channel is about 83% and the surplus channel about 17% (contrasting to 87% and 13% with low initial debt in the text). Thus, the same conclusions are drawn, i.e. the planner uses much more actively the valuation channel in an economy with recursive preferences.

# F Example in an economy with capital

## F.1 Analysis

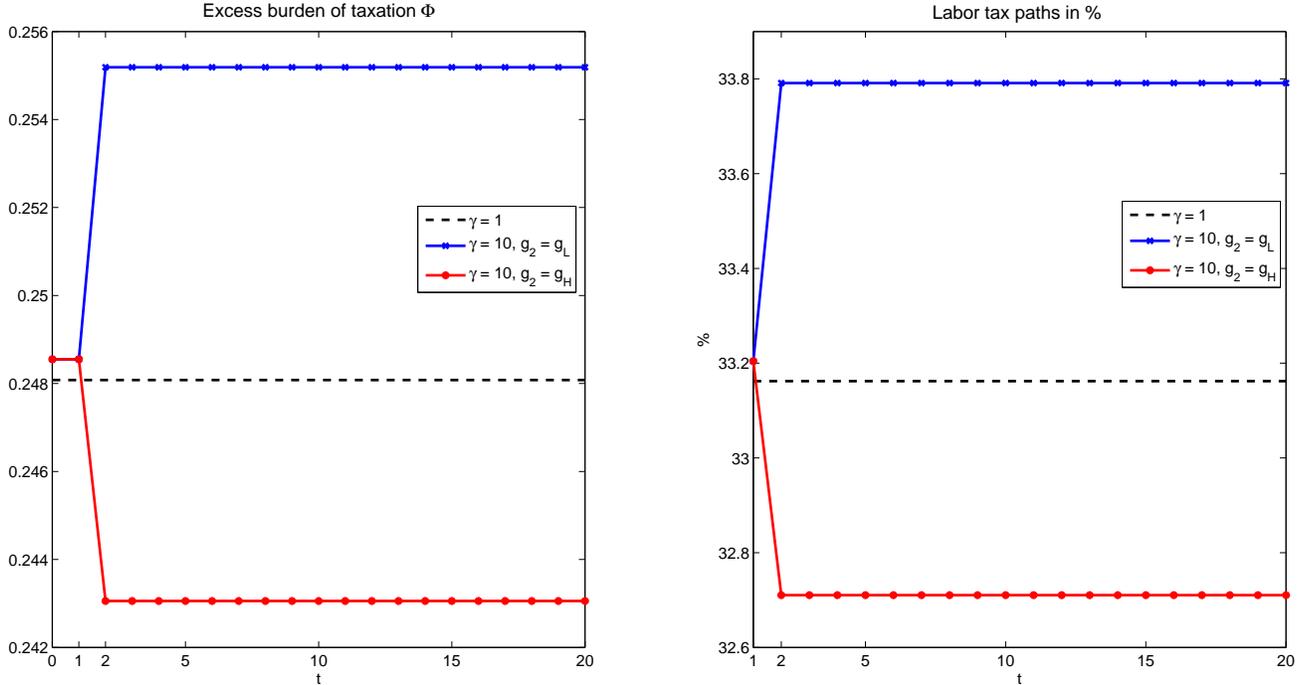


Figure F.1: Paths of the excess burden of taxation and the labor tax for the high-shock and low-shock history.

Consider a simplified stochastic structure – deterministic except for one period. Let government expenditures take two values  $g_L < g_H$ . Assume that government expenditures are low with certainty except for  $t = 2$ . At  $t = 2$  we have  $g_2 = g_H$  with probability  $\pi$  and  $g_2 = g_L$  with probability  $1 - \pi$ . I use superscripts for the endogenous variables in order to denote if we are at the high-shock history ( $g_2 = g_H$ ) or at the low-shock history ( $g_2 = g_L$ ). For example,  $c_t^i, i = H, L$ , denotes consumption at period  $t \geq 2$  when the shock at  $t = 1$  is high or low respectively. Let the utility function be the same as in the numerical exercise section without capital. The calibration is provided in the next section.

The deterministic setup after the second period serves as an example of a case where the excess burden of taxation stays *permanently* at the values it assumes at  $t = 2$ . In particular, since there is no uncertainty before and after  $t = 2$ , we have  $\Phi_1 = \Phi_0$  and  $\Phi_t^i = \Phi_2^i, i = H, L, t \geq 2$ . Turning to the issue of fiscal hedging, we find that the planner is taking a larger wealth position in marginal utility units at the low shock,  $z_2^H < z_2^L$ . As a result, he transfers distortions permanently towards the low-shock history and away from the high-shock history, so  $\Phi_2^H < \Phi_2^L$ . The left panel in figure F.1 plots the respective paths for the excess burden of taxation and the right panel the labor tax dynamics. Recall that this utility function implies a constant labor tax for  $t \geq 1$  in the time-

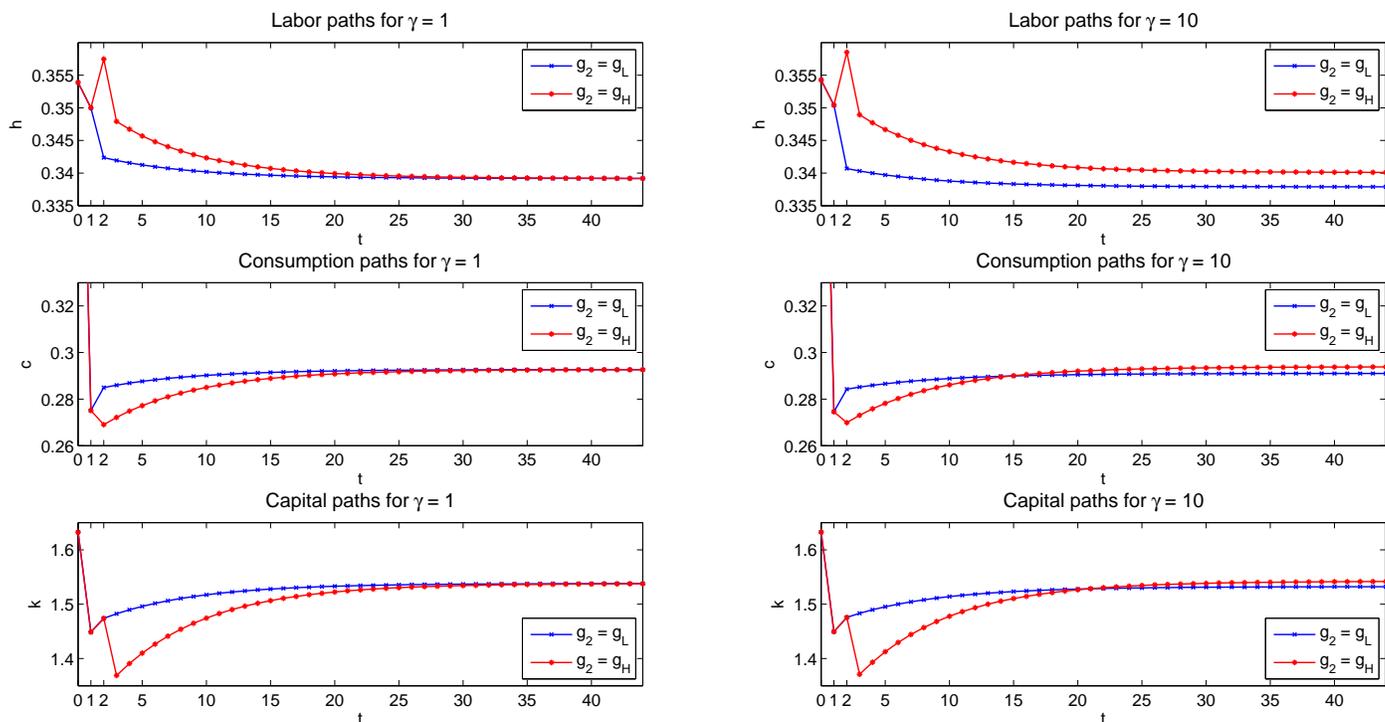


Figure F.2: Left panels depict the labor, consumption and capital paths for the expected utility case ( $\gamma = 1$ ) for the high- and low-shock history. Right panels depict the respective paths for recursive utility ( $\gamma = 10$ ), which converge to two different steady states depending on the realization of the shock at  $t = 2$ .

additive case. With recursive utility, despite the fact that government expenditures revert to a low value with certainty after  $t = 2$ , the labor tax becomes permanently low when there is an adverse shock at  $t = 2$  and permanently high when there is favorable shock at  $t = 2$ .

The planner keeps the the tax rate permanently low or high, because any change in the tax rate in future periods will affect the price of claims at  $t = 2$ , due to the forward-looking nature of continuation utility. Consider for example the high shock-history. If the planner increased the tax rate at any period  $t \geq 3$ , he would decrease the utility of the agent at  $t = 2$ , leading therefore to a higher price of the claim contingent on  $g_2 = g_H$ . This is not optimal though, since the planner is hedging the bad shock with a small position,  $z_2^H < z_2^L$ , and therefore wants to have a low price, i.e. a high state-contingent return on his negative relative wealth position.

Turning to the capital tax, in the time-additive economy there is a *zero* ex-ante capital tax at  $t = 2$  and a zero capital tax for  $t \geq 3$ .<sup>8</sup> For the recursive utility case, the capital tax will be zero

<sup>8</sup>The presence of initial wealth (which would be absent if we had zero initial debt, full depreciation and an initial tax rate on capital income of 100%) alters the taxation incentives for labor income at  $t = 0$  and capital income at  $t = 1$ . In particular, the planner has an incentive to increase initial consumption in order to reduce initial wealth in marginal utility units. By subsidizing initial labor income and taxing capital income at  $t = 1$ , he is able to achieve that. The labor subsidies at the initial period are  $\tau_0 = -17.69\%$  for the time-additive case and  $\tau_0 = -17.76\%$  for the recursive utility case. Following Chari et al. (1994), I do not impose an upper bound on capital taxes. At  $t = 1$  they take the values  $\tau_1^K = 365.31\%$  and  $\tau_1^K = 365.74\%$  for the time-additive and recursive utility case respectively.

for  $t \geq 3$  since the economy becomes deterministic and the utility function belongs to the constant elasticity class. For  $t = 2$ , the ex-ante tax rate will not be zero and its sign depends on the fiscal hedging of the government, as discussed in the text. Figure F.2 plots the time paths for labor, consumption and capital for the two histories. Consumption (labor) at  $t = 2$  is lower (higher) when the expenditure shock is high, putting therefore a larger weight on the state-contingent “subsidy”. As a result, we have an *ex-ante* subsidy, that takes the value of  $-0.5536\%$  in this illustration. In addition, it is worth noting that, since the change in the labor tax is permanent, we have two different steady states depending on what value government expenditures took at  $t = 2$ . For the high-shock history, which is associated with a lower labor tax, the steady state entails higher labor, consumption and capital, whereas for the low-shock history, which is associated with a higher labor tax, the steady state involves lower labor, consumption and capital.

## F.2 Computational details

The production function is  $F = k^\alpha h^{1-\alpha}$ . The parameters for the illustration are  $(\beta, \gamma, \phi_h, \alpha, \delta, \tau_0^K, b_0) = (0.96, 10, 1, 1/3, 0.08, 0.3, 0)$  with a total endowment of time normalized to unity. The parameter  $a_h$  is set so that the household works 0.4 of its time at the first-best steady state. The size of  $g_L$  is set so that the share of government expenditures in the first-best steady state output is 0.22. The high shock is  $g_H = 2 \cdot g_L$  and  $\pi = 0.5$ . The economy features a low shock for each period except for  $t = 2$ , which is the reason why I use a relatively large  $g_H$ .

For the utility function of the example we have  $\Omega(c, h) = 1 - a_h h^{1+\phi_h}$  and  $\tau_t = \tau(\Phi_t) = \Phi_t(1 + \phi_h)/(1 + \Phi_t(1 + \phi_h))$ , which holds only for  $t \geq 1$  due to the presence of initial wealth  $W_0$ . The procedure to solve the problem involves a double loop for the determination of  $\Phi_2^i, i = H, L$  and  $\Phi_0$ .

- *Inner loop*: Fix  $\Phi_0$  and make a guess for  $(\Phi_2^H, \Phi_2^L)$ . Given these two values of the excess burden of taxation, the problem from period  $t = 3$  onward for both histories behaves as a deterministic Ramsey taxation problem, but with different  $\Phi$ 's depending on the high- or low-shock history. In order to solve it, modify the return function as Chari et al. (1994) do, by defining  $\bar{U}(c, 1 - h; \Phi) \equiv U(c, 1 - h) + \Phi\Omega(c, h)$ . For the high-shock history, for  $t \geq 3$  solve the Bellman equation,

$$v^{\text{CCK}}(k) = \max_{c, h, k'} \bar{U}(c, 1 - h; \Phi_2^H) + \beta v^{\text{CCK}}(k')$$

subject to  $c + k' - (1 - \delta)k + g_L = k^\alpha h^{1-\alpha}$ , with the return function  $\bar{U}(c, 1 - h; \Phi_2^H) = \ln c - a_h \frac{h^{1+\phi_h}}{1+\phi_h} + \Phi_2^H(1 - a_h h^{1+\phi_h})$ . For the low-shock history, for  $t \geq 2$ , solve the same Bell-

---

The desire to disentangle the effect of the initial conditions from the effect of uncertainty is the reason why I let the shock materialize at  $t = 2$ .

man equation but with the return function  $\bar{U}(c, 1 - h; \Phi_2^L)$ .

To determine the wealth positions  $z_2^i$  and the respective innovations that allow the update of the guesses for  $\Phi_2^i$ , proceed as follows: Fix  $k_3^H$  and consider the respective Euler equation:

$$\frac{1}{c_2^H} = \beta \frac{1}{c_3^H} \left[ 1 - \delta + \alpha \left( \frac{k_3^H}{h_3^H} \right)^{\alpha-1} \right]$$

Given  $k_3^H$  and the policy functions we found from solving the Bellman equation, the right-hand side is known, determining therefore  $c_2^H$ . Furthermore, use the intratemporal wedge condition for  $g_2 = g_H$  to get  $h_2^H = \left[ \frac{(1-\tau_2^H)(1-\alpha)}{a_h c_2^H} \right]^{\frac{1}{\alpha+\phi_h}} k_2^{\frac{\alpha}{\alpha+\phi_h}}$ , where  $\tau_2^H = \tau(\Phi_2^H)$ . Plug the expression for labor in the resource constraint at  $g_2 = g_H$ ,  $c_2^H + k_3^H - (1-\delta)k_2 + g_H = k_2^\alpha (h_2^H)^{1-\alpha}$  to get one equation in the unknown  $k_2$  and use a non-linear solver to determine it. Furthermore, use the policy functions for  $t \geq 3$  to determine  $v_3^H$  and  $z_3^H$ . Utilities are calculated with the original period utility function (and *not* with the modified  $\bar{U}$ ). Finally, use  $(c_2^H, h_2^H)$  to get  $v_2^H = U(c_2^H, 1 - h_2^H) + \beta v_3^H$  and  $z_2^H = \Omega(c_2^H, h_2^H) + \beta z_3^H$ . Use now the policy functions for the low-shock history to determine  $v_2^L$  and  $z_2^L$  at  $k_2$ . Having the utility values and the wealth positions at  $t = 2$  allows us to calculate the induced likelihood ratios  $m_2^i, i = H, L$ , the market value of the wealth portfolio  $\omega_1 = \pi m_2^H z_2^H + (1 - \pi) m_2^L z_2^L$  and therefore the relative wealth positions  $\eta_2^i = z_2^i - \omega_1, i = H, L$ , given the guess for  $\Phi_2^i$ . Use the innovations  $\eta_2^i$  to update the guess for  $\Phi_2^i$ ,  $\Phi_2^i = \frac{\Phi_0}{1+(1-\beta)(1-\gamma)\eta_2^i \Phi_0}, i = H, L$  and iterate till convergence.

- *Outer loop:* After we reach convergence for  $\Phi_2^i$ , calculate the rest of the allocation for  $t = 0, 1$  given the initial  $\Phi_0$ . In particular, the Euler equation for  $k_2$  is

$$\frac{1}{c_1 \Phi_0} = \beta \pi m_2^H \frac{1}{c_2^H \Phi_2^H} \left[ 1 - \delta + \alpha \left( \frac{k_2}{h_2^H} \right)^{\alpha-1} \right] + \beta (1 - \pi) m_2^L \frac{1}{c_2^L \Phi_2^L} \left[ 1 - \delta + \alpha \left( \frac{k_2}{h_2^L} \right)^{\alpha-1} \right].$$

The right-hand side is known, which delivers  $c_1$ . Express now labor at  $t = 1$  as  $h_1 = \left[ \frac{(1-\tau_1)(1-\alpha)}{a_h c_1} \right]^{\frac{1}{\alpha+\phi_h}} k_1^{\frac{\alpha}{\alpha+\phi_h}}$ ,  $\tau_1 = \tau(\Phi_0)$  and use this expression to solve for  $k_1$  from the resource constraint. Calculate furthermore  $z_1 = \Omega(c_1, h_1) + \beta \omega_1$ . The initial period requires a different treatment due to the presence of initial wealth  $W_0 = b_0 + [(1 - \tau_0^K) \alpha (k_0/h_0)^{\alpha-1} + 1 - \delta] k_0$ . Use the Euler equation for capital to get the initial value of the multiplier  $\lambda_0$ ,  $\lambda_0 = \frac{\beta}{c_1} [1 - \delta + \alpha (k_1/h_1)^{\alpha-1}]$ . Then use the first-order conditions for  $(c_0, h_0)$ , (A.7)-(A.8) and the resource constraint at  $t = 0$  to get a system in three unknowns  $(c_0, h_0, k_0)$  to be solved with a non-linear solver. Update  $\Phi_0$  by calculating the residual in the initial budget constraint,  $I \equiv \Omega(c_0, h_0) + \beta z_1 - \frac{1}{c_0} W_0$ . If  $I > (<) 0$  decrease (increase)  $\Phi_0$  and go back to the inner loop to

redetermine  $\Phi_2^i, i = H, L$  given the new  $\Phi_0$ . Stop when the initial budget constraint holds,  $I = 0$ .

The solution method for the outer loop is based on a fixed value  $k_3^H$ , which delivers in the end an initial value of capital  $k_0$ . I experimented with  $k_3^H$  so that the endogenous initial capital corresponds to 0.9 of the first-best steady state capital. .

There is plethora of methods for solving the Bellman equation. I use the envelope condition method of [Maliar and Maliar \(2013\)](#). I approximate the value function with a 5th degree polynomial in capital and I use 100 grid points. Furthermore, since the steady-state capital depends on  $\Phi_2^i$ , I re-adjust the bounds of the state space for each calculation of the value function in order to focus on the relevant part of the state space. For the high-shock history, I set the lower bound as  $\underline{K} = 0.95 \cdot \min(k_3^H, k_{ss}^H)$  and the upper bound  $\bar{K} = 1.05 \cdot \max(k_3^H, k_{ss}^H)$ . In the same vain, for the low-shock history, I set  $\underline{K} = 0.95 \cdot \min(k_2, k_{ss}^L)$  and  $\bar{K} = 1.05 \cdot \max(k_2, k_{ss}^L)$ . The variables  $k_{ss}^i, i = H, L$  denote the respective steady states.

## G Sequential formulation of Ramsey problem

Readers accustomed to optimal taxation problems with complete markets may wonder how the excess burden of taxation can be *time-varying* when there is a *unique* intertemporal budget constraint. I employ here a sequential formulation of the problem and show the mapping between the optimality conditions of the two formulations in order to make clear where this result is coming from. In short, the time-varying  $\Phi_t$  in the recursive formulation reflects the shadow value of additional “implementability” constraints in the sequential formulation of the problem that arise even in a complete markets setup.<sup>9</sup> The benefit of the recursive formulation of the commitment problem, besides illuminating obviously that  $z$  is the relevant state variable, is the succinct summary of the effects of continuation values in terms of a varying marginal cost of debt. This allows a clean comparison with the time-additive expected utility case. There are obvious similarities in spirit with the optimal risk-sharing literature with recursive preferences, which expresses risk-sharing arrangements in terms of time-varying Pareto weights (see for example [Anderson \(2005\)](#)).<sup>10</sup>

I consider an economy with capital. The specialization of the analysis to an economy without capital is obvious. Let  $X_t \equiv M_t^{\frac{\rho-\gamma}{1-\gamma}}$ ,  $X_0 \equiv 1$ . Let  $v$  refer to the  $\rho$ -transformation of the utility criterion. The Ramsey problem is

$$\max v_0(\{c\}, \{h\})$$

subject to

$$\sum_{t=0}^{\infty} \beta^t \sum_{s^t} \pi_t(s^t) X_t(s^t) \Omega(c_t(s^t), h_t(s^t)) = U_{c_0} W_0 \quad (\text{G.1})$$

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)) \quad (\text{G.2})$$

$$X_{t+1}(s^{t+1}) = m_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} X_t(s^t), \quad X_0 \equiv 1 \quad (\text{G.3})$$

$$v_t(s^t) = U(c_t(s^t), 1 - h_t(s^t)) + \beta \frac{\left[ \sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) \left[ 1 + (1 - \beta)(1 - \rho)v_{t+1}(s^{t+1}) \right]^{\frac{1-\gamma}{1-\rho}} \right]^{\frac{1-\rho}{1-\gamma}} - 1}{(1 - \beta)(1 - \rho)}, \quad t \geq 1 \quad (\text{G.4})$$

---

<sup>9</sup>These constraints describe utility recursions and the law of motion of  $M_t^{\frac{\rho-\gamma}{1-\gamma}}$ . In the case of the multiplier preferences of [Hansen and Sargent \(2001\)](#), it is natural to think of the utility recursions as implementability constraints since they correspond to optimality conditions of the malevolent alter-ego of the household, that minimizes the household’s utility subject to a penalty. See [Karantounias \(2013\)](#). This minimization procedure would also emerge naturally if we expressed recursive utility as the variational utility of [Geoffard \(1996\)](#).

<sup>10</sup>Note also that recursive utility adds  $z$  as a state variable to the optimal taxation problem, whereas  $z$  can be ignored in the time-additive case. The reason is that  $z$  is necessary for the determination of the Ramsey plan only though its shadow cost,  $\Phi$ . When the excess burden of taxation is constant, the return function of the second-best problem can be augmented in such a way, so that  $z$  becomes redundant as a state variable. See [Lucas and Stokey \(1983\)](#) or [Zhu \(1992\)](#) and [Chari et al. \(1994\)](#).

where  $W_0 \equiv R_0^K k_0 + b_0$ ,  $(b_0, k_0, s_0, \tau_0^K)$  given, and  $m_{t+1} = \frac{[1+(1-\beta)(1-\rho)v_{t+1}]^{\frac{1-\gamma}{1-\rho}}}{E_t[1+(1-\beta)(1-\rho)v_{t+1}]^{\frac{1-\gamma}{1-\rho}}}$ .

Assign multipliers  $\bar{\Phi}$ ,  $\beta^t \pi_t \lambda_t$ ,  $\beta^t \pi_t \nu_t$  and  $\beta^t \pi_t \xi_t$  on (G.1), (G.2), (G.3) and (G.4) respectively. The derivative of the utility index with respect to  $c_{t+i}$  can be calculated recursively from the relationship  $\frac{\partial V_t}{\partial c_{t+i}} = \frac{\partial V_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial V_{t+1}} \frac{\partial V_{t+1}}{\partial c_{t+i}}$ ,  $i \geq 1$ . Similarly for labor. This leads to  $\frac{\partial V_0}{\partial c_t} = (1-\beta)V_0^\rho \beta^t \pi_t X_t U_{ct}$  and  $\frac{\partial V_0}{\partial h_t} = -(1-\beta)V_0^\rho \beta^t \pi_t X_t U_{lt}$ . For the  $\rho$ -transformation that we use here we have  $\frac{\partial v_0}{\partial c_t} = \beta^t \pi_t X_t U_{ct}$  and  $\frac{\partial v_0}{\partial h_t} = -\beta^t \pi_t X_t U_{lt}$ . The first-order necessary conditions are

$$c_t, t \geq 1: \quad X_t(s^t)U_c(s^t) + \bar{\Phi}X_t(s^t)\Omega_c(s^t) + \xi_t(s^t)U_c(s^t) = \lambda_t(s^t) \quad (\text{G.5})$$

$$h_t, t \geq 1: \quad -X_t(s^t)U_l(s^t) + \bar{\Phi}X_t(s^t)\Omega_h(s^t) - \xi_t(s^t)U_l(s^t) = -\lambda_t(s^t)F_H(s^t) \quad (\text{G.6})$$

$$k_{t+1}(s^t), t \geq 0: \quad \lambda_t(s^t) = \beta \sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) \lambda_{t+1}(s^{t+1}) [1 - \delta + F_K(s^{t+1})] \quad (\text{G.7})$$

$$X_t(s^t), t \geq 1: \quad \nu_t(s^t) = \bar{\Phi}\Omega_t(s^t) + \beta \sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) m_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} \nu_{t+1}(s^{t+1}) \quad (\text{G.8})$$

$$v_t(s^t), t \geq 1: \quad \xi_t(s^t) = (1-\beta)(\rho-\gamma)X_t(s^t)\phi_t(s^t) + m_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \xi_{t-1}(s^{t-1}), \quad (\text{G.9})$$

where

$$\phi_t(s^t) \equiv V_t(s^t)^{\rho-1} \nu_t(s^t) - \mu_t(s^t)^{\rho-1} \sum_{s_t} \pi_t(s_t|s^{t-1}) m_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \nu_t(s^t),$$

and  $\xi_0 \equiv 0$ . The optimality conditions with respect to the initial consumption-labor allocation are (A.7) and (A.8).

I will show now the mapping between the sequential formulation and the recursive formulation and in particular the relationship between the time-varying  $\Phi_t$  and  $\xi_t$ . Solve at first (G.8) forward to get

$$\nu_t = \bar{\Phi} E_t \sum_{i=0}^{\infty} \beta^i \frac{X_{t+i}}{X_t} \Omega_{t+i}$$

and therefore  $\nu_t = \bar{\Phi} U_{ct} W_t = \bar{\Phi} z_t$ , i.e.  $\nu_t$  – the shadow value to the planner of an increase in  $X_t$  – is equal to wealth (in marginal utility terms) times the cost of taxation  $\bar{\Phi}$ . Thus,  $\phi_t$  – the “innovation” in the multiplier  $\nu_t$  – is equal to a multiple of  $\eta_t$ ,  $\phi_t = \bar{\Phi} \eta_t$ . Furthermore, define the scaled multiplier  $\tilde{\xi}_t \equiv \xi_t / X_t$ ,  $\tilde{\xi}_0 \equiv 0$  and note that it follows the law of motion

$$\begin{aligned}
\tilde{\xi}_t &= (1 - \beta)(\rho - \gamma)\phi_t + \tilde{\xi}_{t-1} \\
&= (1 - \beta)(\rho - \gamma) \sum_{i=1}^t \phi_i = (1 - \beta)(\rho - \gamma) \sum_{i=1}^t \eta_i \bar{\Phi}
\end{aligned}$$

Turn now to the multiplier in the text which, when solved backwards, delivers  $\Phi_t = \Phi_0 / (1 + (1 - \beta)(\rho - \gamma) \sum_{i=1}^t \eta_i \Phi_0)$ , where  $\Phi_0$  is the multiplier on the initial period implementability constraint. Thus, by setting  $\Phi_0 = \bar{\Phi}$  we have

$$\Phi_t = \frac{\bar{\Phi}}{1 + \tilde{\xi}_t}, \quad (\text{G.10})$$

or, in terms of the non-scaled  $\xi_t$ ,  $\Phi_t = \bar{\Phi} X_t / (X_t + \xi_t)$ . Therefore, the time-varying excess burden of taxation captures the shadow value of continuation utilities that determine intertemporal marginal rates of substitution. Consider now the multipliers  $\lambda_t$  in the sequential formulation and their relationship to their counterparts in the recursive formulation,  $\lambda_t^R$ . (G.5) can be written as  $U_{ct} + \frac{\bar{\Phi} X_t}{X_t + \xi_t} \Omega_{ct} = \frac{\lambda_t}{X_t + \xi_t}$ . Recall that the optimality condition with respect to consumption in the recursive formulation is  $U_{ct} + \Phi_t \Omega_{ct} = \lambda_t^R$ . Use then (G.10) to get that  $\lambda_t = (X_t + \xi_t) \lambda_t^R$ . Thus,

$$\begin{aligned}
\frac{\lambda_{t+1}}{\lambda_t} &= \frac{X_{t+1} + \xi_{t+1}}{X_t + \xi_t} \frac{\lambda_{t+1}^R}{\lambda_t^R} = \frac{X_{t+1}}{X_t} \frac{\frac{\bar{\Phi} X_t}{X_t + \xi_t}}{\frac{\bar{\Phi} X_{t+1}}{X_{t+1} + \xi_{t+1}}} \frac{\lambda_{t+1}^R}{\lambda_t^R} \\
&= m_{t+1}^{\frac{\rho - \gamma}{1 - \gamma}} \frac{\lambda_{t+1}^R}{\lambda_t^R} \frac{\Phi_t}{\Phi_{t+1}}.
\end{aligned}$$

Thus, (G.7) delivers the same condition as equation  $E_t S_{t+1}^* (1 - \delta + F_{K,t+1}) = 1$  in the text.

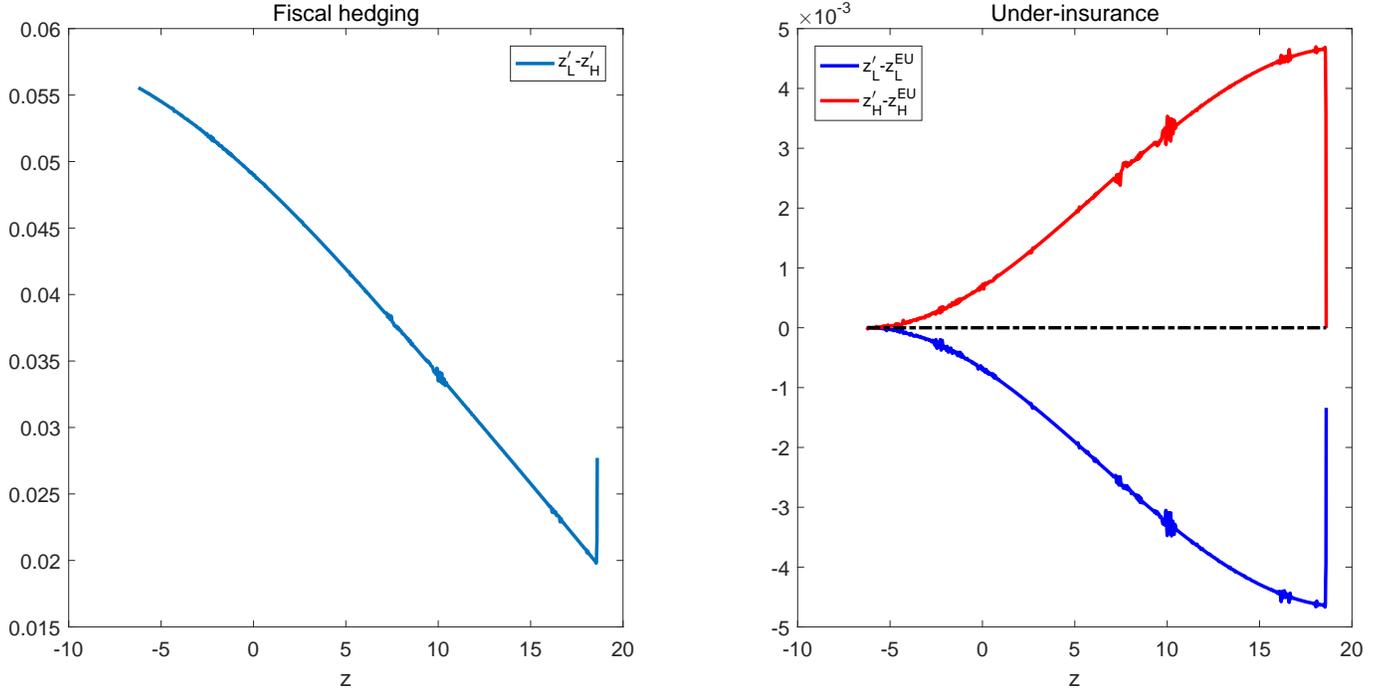


Figure H.1: The graph denotes the fiscal hedging the the “under-insurance” in comparison to expected utility, when there is preference for late resolution of uncertainty.

## H Preference for late resolution of uncertainty

Assume the same utility function as in the numerical exercise in the main text and set  $\gamma = 0 < \rho = 1$ . Let the rest of the preference parameters and the shocks be the same as in the baseline exercise with no persistence. The left graph of figure H.1 shows that the planner still hedges adverse shocks by selling more claims against good times and less claims against bad times. But debt becomes now more expensive in good times, due to the love of future utility volatility. Therefore, the planner issues less debt against good times (and taxes less) and more debt against bad times (and taxes more), effectively “under-insuring” with respect to expected utility. This is displayed in the right graph of figure H.1. Table H.1 displays the positive correlation of changes in tax rates with government spending and figure H.2 displays the ensemble moments of the tax rate and the debt ratio. Note that the positive drift is very small for this parametrization.

In the computational section I highlighted the non-convexities in the implementability constraint budget that emerge with recursive utility. These non-convexities disappear in the case of  $\rho > \gamma$  as figure H.3 shows. This allows also the increase of the upper bounds of the state space. I build the state space with  $\bar{\Phi} = 3$ , which corresponds to a tax rate of 85.41% and to upper bounds  $(\bar{z}_L, \bar{z}_H) = (18.6064, 18.5773)$ . These upper corresponds to values of debt that are 10.99 and 9.8 multiples of output. In (very) long simulations we noted that the upper bound is not innocuous, in the sense that the tax rate tends to put most of its mass towards it, as can be seen in figure 4(a). Figure 4(b) plots the stationary distributions of the tax rate and debt, which exhibit large

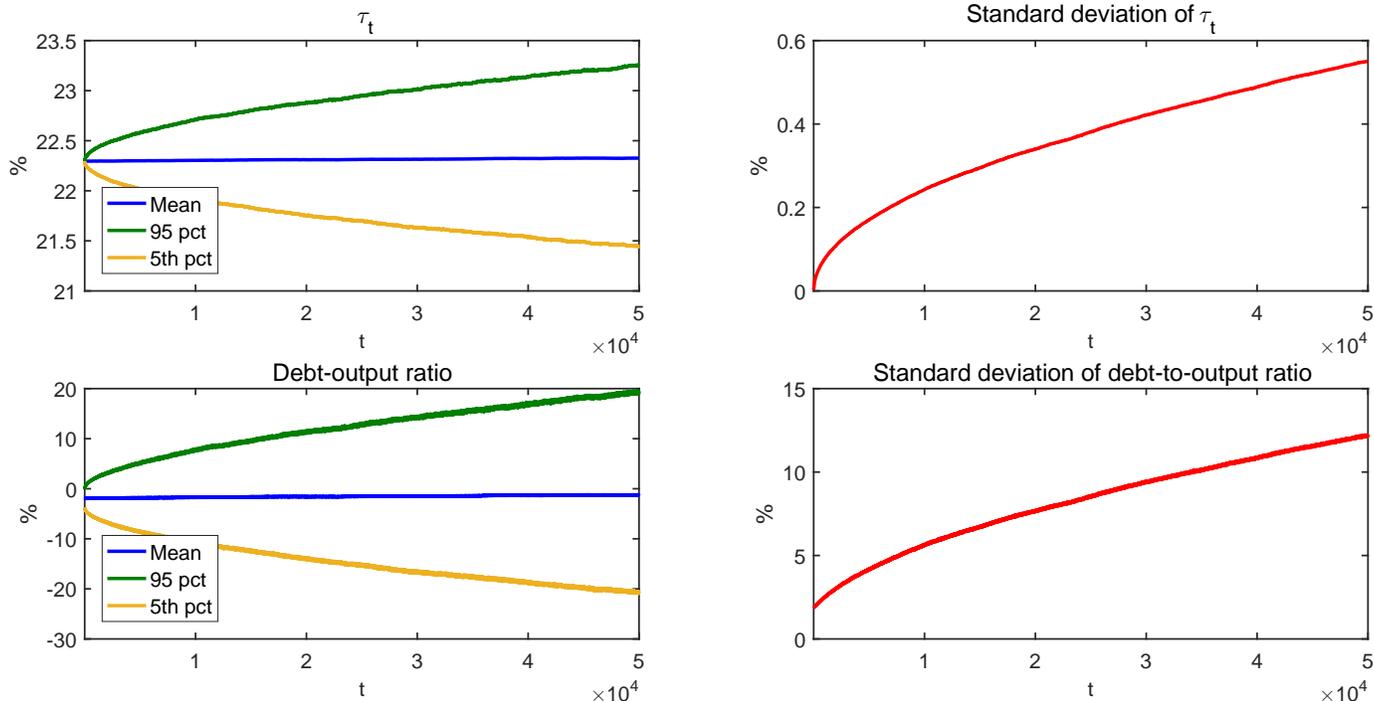


Figure H.2: Ensemble moments of 10,000 sample paths of 50,000 period length. The increase in the mean tax rate is very small for sample paths of this length and the particular calibration.

Table H.1: Statistics of sample paths for late resolution of uncertainty.

	<i>Recursive utility</i>		
	200 periods	2000 periods	50000 periods
Autocorrelation of $\tau$	0.9792	0.9979	0.9999
Correlation of $\Delta\tau$ with $g$	1	1	0.9999
Correlation of $\Delta\tau$ with output	1	0.9999	0.9963
Correlation of $\Delta b$ with $g$	-0.6991	-0.6974	-0.6968
Correlation of $\Delta b$ with $\Delta\tau$	-0.6991	-0.6974	-0.6971
Correlation of $\tau$ with $g$	0.1105	0.0353	0.0070
Correlation of $\tau$ with output	0.1049	0.0177	-0.0775
Correlation of $b$ with $\tau$	0.0302	0.3656	0.9049

The table reports median sample statistics across 10000 sample paths of variable lengths.

probability mass at the right tails (the debt distribution is actually bimodal). This is in contrast to the baseline exercise of the paper with aversion to utility volatility, where the upper tails are thin.

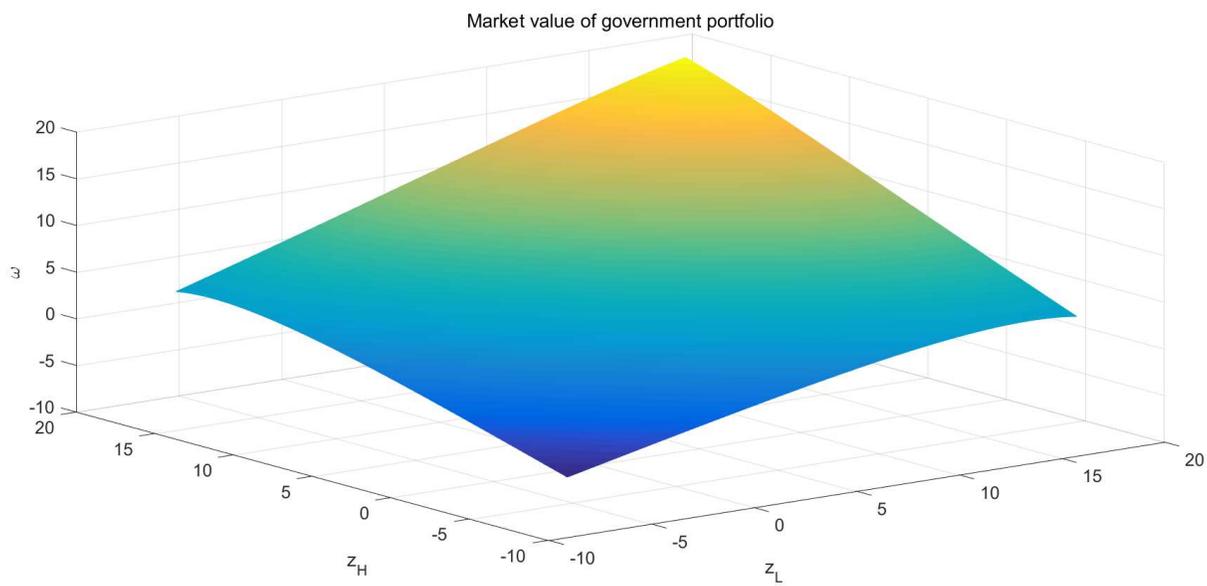
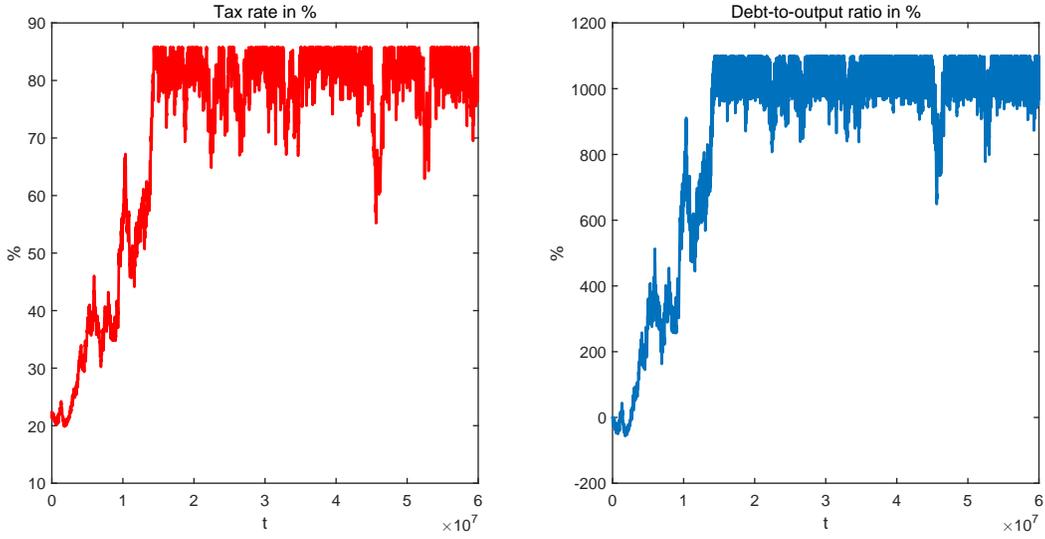
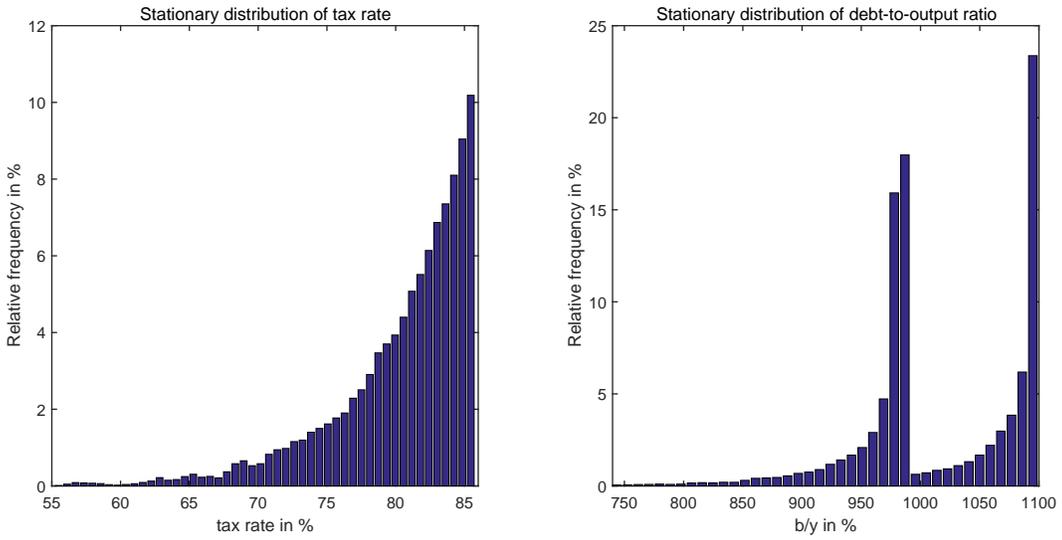


Figure H.3: The market value of debt  $\omega$  as function of the state-contingent positions when  $\gamma = 0 < \rho = 1$ .



(a) Long Simulation.



(b) Stationary distributions.

Figure H.4: Stationary distributions from a simulation of 60 million periods. The first 20 million periods were dropped. The first and second moments (in %) are  $(E(\tau), \sqrt{\text{Var}(\tau)}) = (80.38, 4.98)$  and  $(E(b/y), \sqrt{\text{Var}(b/y)}) = (1014.97, 69.46)$ .

## 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.
- Anderson, Evan W. 2005. The dynamics of risk-sensitive allocations. *Journal of Economic Theory* 125 (2):93–150.
- Benigno, Pierpaolo and Michael Woodford. 2006. Optimal taxation in an RBC model: A linear-quadratic approach. *Journal of Economic Dynamics and Control* 30 (9-10):1445–1489.
- Berndt, Antje, Hanno Lustig, and Sevin Yeltekin. 2012. How Does the U.S. Government Finance Fiscal Shocks? *American Economic Journal: Macroeconomics* 4:69–104.
- 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.
- Christiano, Lawrence J. and Martin Eichenbaum. 1992. Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations. *American Economic Review* 82 (3):430–450.
- Farhi, Emmanuel. 2010. Capital Taxation and Ownership when Markets are Incomplete. *Journal of Political Economy* 118 (5):908–948.
- Geoffard, Pierre-Yves. 1996. Discounting and optimizing: Capital accumulation problems as variational minmax problems. *Journal of Economic Theory* 69:53–70.
- Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.
- Karantounias, Anastasios G. 2013. Managing pessimistic expectations and fiscal policy. *Theoretical Economics* 8 (1):193–231.
- Lucas, Robert Jr. and Nancy L. Stokey. 1983. Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics* 12 (1):55–93.
- Maliar, Lilia and Serguei Maliar. 2013. Envelope Condition Method versus Endogenous Grid Method for Solving Dynamic Programming Problems. *Economic Letters* 120:262–266.
- Martin, Ian. 2012. On the Valuation of Long-Dated Assets. *Journal of Political Economy* 120 (2):346–358.
- Zhu, Xiaodong. 1992. Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58 (2):250–289.