FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Fiscal Austerity in Ambiguous Times

Axelle Ferrière and Anastasios G. Karantounias

Working Paper 2016-6a
March 2016 (revised January 2018)

Abstract: This paper analyzes optimal fiscal policy with ambiguity aversion and endogenous government spending. We show that, without ambiguity, optimal surplus-to-output ratios are acyclical and that there is no rationale for either reduction or further accumulation of public debt. In contrast, ambiguity about the cycle can generate optimally policies that resemble "austerity" measures. Optimal policy prescribes higher taxes in adverse times and front-loaded fiscal consolidations that lead to a balanced primary budget in the long-run. This is the case when interest rates are sufficiently responsive to cyclical shocks—that is, when the intertemporal elasticity of substitution is sufficiently low.

JEL classification: D80, E62, H21, H63

Key words: public consumption, intertemporal elasticity of substitution, balanced budget, austerity, fiscal consolidation, ambiguity aversion, multiplier preferences

The authors thank David Backus, Richard Rogerson, Thomas J. Sargent, Stanley E. Zin for their support of this project and Mikhail Golosov for his discussion. They are grateful to the editor and to three anonymous referees for insightful comments. They also thank Roc Armenter, Anmol Bhandari, Lukasz Drozd, Murat Tasci, conference participants at the Ambiguity and Robustness Workshops at the Becker Friedman Institute and New York University, the CEF Meetings in Oslo, the 15th Conference on Research on Economic Theory and Econometrics at Tinos, the EEA Meetings in Toulouse, the SED Meetings in Toronto, and seminar participants at the Federal Reserve Banks of Atlanta, Cleveland, and Philadelphia and at Florida International University and Florida State University. 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 Axelle Ferrière, Department of Economics, European University Institute, Via delle Fontanelle 18, 50014, Fiesole, Italy, axelle.ferriere@eui.eu, or Anastasios G. Karantounias (corresponding author), Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309-4470, 404-498-8825, anastasios.karantounias@atl.frb.org.

1 Introduction

This paper studies the optimal determination of government spending, taxes and debt in an environment of *ambiguity* about the cycle. We uncover the central role of the intertemporal elasticity of substitution (IES) for fiscal policy under ambiguity. Our main finding is that "austerity" can become optimal in such an economy if interest rates are sufficiently responsive to cyclical shocks, that is, when the IES is below unity. Optimal policy prescribes then front-loaded fiscal consolidations and convergence to a balanced primary budget in the long-run.

We take seriously the idea that economic agents face uncertainty about the cycle that cannot be specified with a unique probability measure, and are *averse* towards it. In our analysis we treat all margins of fiscal policy as equally important, which is why we endogenize government consumption by allowing it provide utility. Our study is relevant for answering questions about the optimal fiscal *mix* and the optimal debt management under ambiguity aversion.

Our environment features an economy without capital and complete markets as in Lucas and Stokey (1983). Our government chooses distortionary labor taxes, government consumption and issues state-contingent debt to maximize the utility of the representative household. To introduce ambiguity about the cycle, we assume doubts about the probability model of technology shocks. We use the multiplier preferences of Hansen and Sargent (2001) to capture our household's aversion towards this ambiguity.

As a first step, we analyze optimal fiscal policy without ambiguity. To capture distortions at the government consumption margin, we define a new wedge at the second-best, the *public wedge*. Our basic finding in a setup with full confidence in the model is that the optimal allocation, public wedge and tax rate are history-independent, extending the Lucas and Stokey result to environments with utility-providing government spending. Using a standard homothetic specification for the utility of private and government consumption and assuming a constant Frisch elasticity of labor supply furnishes a comprehensive *smoothing* result: *both* the share of government consumption in output *and* taxes are constant. Optimal policy prescribes a deficit at the initial period and an *acyclical* surplus-to-output ratio afterwards. Public debt remains stationary, without exhibiting negative or positive drifts. Consequently, *neither* fiscal consolidations, *nor* further accumulation of public debt are optimal.

There are stark differences when we turn to the analysis of the optimal fiscal policy in an environment with ambiguity. The planner still runs a deficit at the initial period but *both* the subsequent acyclicality of distortions *and* the lack of drifts in public debt break down. We find that two, diametrically opposite, policies can be optimal, depending on the size of the IES relative to unity: when the IES is below unity and equilibrium interest rates are very responsive to changes in consumption, we find that countercyclical tax rates are optimal, i.e. taxes increase in bad times and decrease in good times.[1] Furthermore, it is also optimal to *reduce* on average public debt and

[1]The terms countercyclical and procyclical refer respectively to negative or positive correlation with output

tax rates till debt becomes zero and a balanced primary budget is reached. These two facets of optimal policy is what we call "austerity" policy. In contrast, the opposite, "anti-austerity," policy emerges with an IES larger than unity. Tax rates increase in good times and decrease in bad times. Furthermore, the procyclicality of tax distortions is coupled with increasing –on average– public debt and taxes over time.

The main mechanism in an environment with ambiguity is based on the *endogenous* pessimistic beliefs of the household, which alter in a non-trivial way the optimal policy problem. In particular, a cautious household assigns high probability on low utility events. The household's utility, and therefore its probability assessments, depend though on policy variables. A Ramsey planner recognizes this dependence, and by setting taxes, manages the pessimistic expectations of the household. In particular, high future taxes, by reducing the utility of the household, raise the pessimistic probabilities and therefore *increase* equilibrium prices of state-contingent claims, reducing therefore the return of state-contingent debt. Similarly, low future taxes decrease equilibrium prices of state contingent debt and increase the return on debt.

How does the government manage this endogenous pessimism? The government uses the pessimistic beliefs of the household in order to *amplify* the present discounted value of surpluses. This type of policy takes a very intuitive form when the government issues a *portfolio* of state-contingent debt: the government increases –by taxing more– the pessimistic probability weight on high "values" of surpluses and reduces –by taxing less– the weight on low "values" of surpluses. By "value" of surpluses we mean surpluses multiplied by the *marginal utility* of consumption.[2] Such a policy increases the total revenue from debt issuance, relaxing therefore the government budget and the need for distortionary taxation.

Why does the IES enter the discussion? The response of the "value" of surpluses to shocks, and therefore the increase or decrease in taxes, depends obviously on the elasticity of marginal utility, and thus on the IES. To see that, consider an IES that is smaller than unity and assume a negative productivity shock. Surpluses fall due to a reduction in output. But a contraction of output, and therefore of consumption, leads to an *expansion* of marginal utility, making therefore the behavior of their product – the "value" of surpluses – ambiguous. When the household does not substitute a lot intertemporally, marginal utility is very elastic, inducing therefore a big increase in the intertemporal rate of substitution and therefore, a big drop in the state-contingent return of debt. Thus, when the IES is lower than unity, the decrease in the return in bad times (by means of marginal utility) over-compensates the decrease in surpluses, leading to an *increase* in the "value" of surpluses. Consequently, by issuing more state-contingent debt and taxing more against adverse times, the planner *amplifies* –through the pessimistic beliefs– the decrease in recessionary interest rates and raises additional revenue.

throughout the paper.

[2] To calculate the presented discounted value of surpluses, the planner needs to take into account both pessimistic beliefs and marginal utilities.

The opposite policy is followed when the IES is larger than unity. Marginal utility is not responsive enough, and thus the "value" of surpluses in good times remains higher than the "value" of surpluses in bad times. Thus, the planner will again *amplify* these values through the pessimistic beliefs by using now a procyclical tax rate. In the knife-edge case of a unitary IES, values of surpluses are constant across the cycle, muting therefore the incentives for active expectation management and leading to the same fiscal policies as without ambiguity. To conclude, by controlling the sensitivity of interest rates to consumption growth, the IES acquires a novel role with ambiguity aversion: it indicates which states of the world have the most potential for interest rate manipulation through the channel of the endogenous pessimistic beliefs.

Lastly, if we take the stance that doubts about the model are unfounded, i.e. if we assume that the probability model that the agents distrust is actually the true data-generating process, then the IES determines also the long-run results about the drift in taxes and public debt. To see that, assume that the IES is below unity. Then good times bear low taxes. But good times happen more often according to the true model than what the pessimistic household expects. Thus, low-tax events happen relatively often, which leads to a decrease in taxes and debt over time till the balanced budget is reached, a point where price manipulation becomes irrelevant since public debt is zero. The opposite is true in the high IES, "anti-austerity" case. Good times are associated with high taxes, and since they happen relatively often, we have an actual increase of taxes and debt over time.

1.1 Related literature

Optimal taxation studies typically treat government expenditures as exogenous, abstracting from questions about the optimal *mix* of taxes and spending. Teles (2011) raises valid concerns about this practice, by showing that the exogenous specification of the level or share of government consumption can alter non-trivially both the interpretation and the welfare consequences of optimal policy. The positive study of Bachmann and Bai (2013) is a notable exception: they endogenize spending and build a business cycle model that successfully captures the basic cyclical features of public consumption. Their setup involves though a balanced budget, and is therefore not useful for answering questions about public debt.[3] Klein et al. (2008) and Debortoli and Nunes (2013) explore optimal taxation with endogenous spending in a deterministic setup and drop the commitment assumption.

Our paper is also related to the literature on fiscal consolidations. Taking as given their necessity, Romei (2014) studies the effects of debt reduction in a heterogenous agents economy, whereas Bi et al. (2013) focus on the uncertainty that may surround the timing and composition of consolidation measures. In contrast, Dovis et al. (2016) have studied how the interaction of

[3] For an early study in the same vein, see Ambler and Paquet (1996). See also Stockman (2001) for the welfare analysis of balanced-budget rules and Kydland and Prescott (1980) for an early contribution.

inequality and lack of commitment can *optimally* lead to cycles between austerity and populistic regimes.[4]

Several papers have modeled ambiguity aversion by using the multiplier preferences of Hansen and Sargent (2001). For example, Bidder and Smith (2012) focused on environments with nominal frictions, Benigno and Nisticò (2012) on international portfolio choice, Pouzo and Presno (2016) on sovereign default, whereas Croce et al. (2012) studied the effects of technological and fiscal uncertainty on long-run growth.[5]

We follow a smooth approach to ambiguity aversion. Nonetheless, of particular interest is the work of Ilut and Schneider (2014), who show that confidence shocks can be a substantial driver of fluctuations at the labor margin. Setups with kinks have also non-trivial implications for endogenous asset supply (Bianchi et al. (2017)) and can also lead to interesting inertia in price-setting (Ilut et al. (2016)).[6]

Fears of model misspecification feature also in the fiscal policy analysis of Karantounias (2013a) and in the monetary policy analysis of Benigno and Paciello (2014), Barlevy (2009) and Barlevy (2011). In Karantounias (2013a), the management of the household's pessimistic expectations played a prominent role. However, government expenditures were treated as exogenous. Furthermore, the analysis was based on paternalism: the policymaker had full confidence in the model, whereas the household did not. Here instead, we use a planner that adopts the perspective of the household in evaluating welfare and proceed also to the numerical evaluation of optimal policy.

This paper uses recursive methods developed in Karantounias (2013b), who provides a comprehensive analysis of optimal labor and capital taxation with recursive preferences in the typical setup of exogenous government expenditures. The connection with the current work comes from the fact that both recursive utility – if we assume preference for early resolution of uncertainty– and multiplier preferences, imply –for different reasons– effectively *aversion* to volatility in continuation utilities, and therefore, lead to a similar mechanism of pricing kernel manipulation.[7] The same would be generally true for *any* kind of preferences that result in aversion to volatility in continuation utilities.

The crucial difference in the current setup though is the endogenous government consumption margin, a feature which may lead to surprising results even for a unitary IES, which is the case where these two classes of preferences are observationally equivalent. For example, in the current paper we prove that optimal policy is the same as without ambiguity when we have unitary IES, whereas Karantounias (2013b) demonstrates that optimal policy is significantly different from the case where time and risk attitudes are not disentangled, even for unitary IES.

[4]There is a large empirical literature that looks at fiscal adjustments. See for example the seminal contribution of Alesina and Perotti (1995).

[5]Of independent interest is also the work of Boyd (1990), who studies the optimal deterministic growth problem with the recursive preferences of Koopmans (1960).

[6]See Epstein and Schneider (2010) for a survey of the implications of ambiguity aversion for asset prices.

[7]We invite the reader to entertain this alternative interpretation, by calculating timing premia as in Epstein et al. (2014).

The difference is coming from the way the "value" of surpluses reacts to shocks when government spending is exogenous. We show analytically in the last section of the paper that the sharp role of the size of the IES relative to unity is absent when government spending is exogenous, because the sensitivity of the "value" of surpluses depends also on other parameters, like the level of taxes and the share of government spending. Furthermore, we explore quantitatively the implications of exogenous spending in environments with ambiguity aversion, and find that the "values" of surpluses are typically procyclical, even when the IES is low. Thus, the "anti-austerity" result is typically optimal when government spending is exogenous. On a more general level, our analysis demonstrates that the question of optimal financing of government spending under ambiguity aversion cannot be independent of what these expenditures do in the economy and how they affect interest rates.

1.2 Organization

Section 2 describes the economy with full confidence in the model and section 3 sets up the Ramsey problem with utility-providing government consumption and derives the properties of optimal policy. Section 4 describes an economy with doubts about the probability model of technology shocks and displays the problem of a planner that adopts the welfare criterion of the household. Section 5 analyzes the basic cyclicality and drift properties of optimal policy with ambiguity aversion and highlights the prominent role of the IES. Section 6 performs numerical exercises and section 7 contrasts our optimal plan to the case where government consumption is exogenous. Section 8 concludes. The Appendix provides proofs of propositions and details of the numerical method. A separate Online Appendix contains details about our expansion around the balanced budget that may be of independent interest.

2 Economy

Time is discrete and the horizon is infinite. We use a complete markets economy without capital as Lucas and Stokey (1983). Government expenditures are endogenous and provide utility to the representative household. Let s_t denote the technology shock at time t and let $s^t \equiv (s_0, s_1, ..., s_t)$ denote the partial history of shocks up to period t with probability $\pi_t(s^t)$. There is no uncertainty at $t = 0$, so $\pi_0(s_0) \equiv 1$. The operator E denotes expectation with respect to π throughout the paper. The resource constraint of the economy reads

$$c_t(s^t) + g_t(s^t) = s_t h_t(s^t), \tag{1}$$

where $c_t(s^t)$ private consumption, $g_t(s^t)$ government consumption and $h_t(s^t)$ labor. The notation indicates the measurability of these functions with respect to the partial history s^t. Total endowment of time is normalized to unity, so leisure is $l_t(s^t) = 1 - h_t(s^t)$.

Household. The representative household derives utility from stochastic streams of private consumption, leisure and government consumption. Its preferences are

$$\sum_{t=0} \beta^t \sum_{s^t} \pi_t(s^t) U(c_t(s^t), 1 - h_t(s^t), g_t(s^t)) \tag{2}$$

where U is monotonic and concave. The household works at the pre-tax wage $w_t(s^t)$, pays proportional taxes on its labor income with rate $\tau_t(s^t)$ and trades in complete asset markets. Let $b_{t+1}(s^{t+1})$ denote the holdings of an Arrow security that promises one unit of consumption if the state of the world is s_{t+1} next period and zero otherwise. This security trades at the price of $p_t(s_{t+1}, s^t)$ in units of consumption at history s^t.

In order to ease notation, let $x \equiv \{x_t(s^t)\}_{t,s^t}$ stand for an arbitrary stochastic process x. Given prices (p, w) and government policies (τ, g), the household chooses $\{c, h, b\}$ to maximize (2) subject to

$$c_t(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}) \leq (1 - \tau_t(s^t)) w_t(s^t) h_t(s^t) + b_t(s^t), \tag{3}$$

and the constraints $c_t(s^t) \geq 0, h_t(s^t) \in [0, 1]$, where b_0 is given. The household is also subject to the no-Ponzi-game condition

$$\lim_{t\to\infty} \sum_{s^{t+1}} q_{t+1}(s^{t+1}) b_{t+1}(s^{t+1}) \geq 0 \tag{4}$$

where $q_t(s^t) \equiv \prod_{j=0}^{t-1} p_j(s_{j+1}, s^j)$ denotes the price of an Arrow-Debreu contract at $t = 0$ with the normalization $q_0 \equiv 1$.

A representative competitive firm operates the linear technology. The government chooses spending, collects tax revenues and trades with the household in Arrow securities. The government budget constraint reads

$$b_t(s^t) = \tau_t(s^t) w_t(s^t) h_t(s^t) - g_t(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}). \tag{5}$$

Competitive equilibrium. A competitive equilibrium is a collection of prices (p, w), a private consumption-labor allocation (c, h), Arrow securities holdings b and government policies (τ, g) such that 1) given (p, w) and (τ, g), (c, h, b) solves the household's problem, 2) given w firms maximize profits, 3) prices (p, w) are such so that markets clear, i.e. the resource constraint (1) holds.[8]

2.1 Optimality conditions

Profit maximization of the competitive firm equates the wage to the marginal product of labor, $w_t = s_t$. Taking government consumption as exogenously given, the household supplies labor according to

$$\frac{U_l(c_t, 1 - h_t, g_t)}{U_c(c_t, 1 - h_t, g_t)} = (1 - \tau_t)w_t, \tag{6}$$

which equates the marginal rate of substitution of consumption and leisure with the after-tax wage. The optimal decision with respect to Arrow securities is characterized by

$$p_t(s_{t+1}, s^t) = \beta\pi_{t+1}(s_{t+1}, s^t)\frac{U_c(s^{t+1})}{U_c(s^t)}, \tag{7}$$

which equates the marginal rate of substitution of consumption at s^{t+1} for consumption at s^t with the price of an Arrow security. The respective price of an Arrow-Debreu contract at $t = 0$ is $q_t(s^t) = \beta^t\pi_t(s^t)\frac{U_c(s^t)}{U_c(s_0)}$. Note furthermore that the asymptotic condition (4) holds in equilibrium with equality, which leads to the exhaustion of the household's unique intertemporal budget constraint.

3 Ramsey problem with full confidence in the model

Consider the problem of the Ramsey planner that chooses under commitment at $t = 0$ government expenditures, distortionary taxes and state-contingent debt in order to maximize the utility of the representative household at the competitive equilibrium. Before we proceed to this problem, it is instructive to understand the first-best allocation, i.e. the allocation that could be sustained as a competitive equilibrium if lump-sum taxes were available.

[8] Note that we have not used a separate notation b_t^g for the government's asset holdings but have instead used the fact that in equilibrium $b_t^g = -b_t$.

3.1 First-best problem

The first-best problem is to choose the allocation $c_t, g_t \geq 0, h_t \in [0,1]$ in order to maximize the utility of the representative household (2) subject to the resource constraint of the economy (1). The optimal allocation is characterized by the resource constraint and two optimality conditions,

$$\frac{U_g(c,1-h,g)}{U_c(c,1-h,g)} = 1 \tag{8}$$

$$\frac{U_l(c,1-h,g)}{U_c(c,1-h,g)} = s. \tag{9}$$

Equation (8) equates the marginal rate of substitution of government for private consumption with the respective marginal rate of transformation, which is unity. Thus, the first-best provision of government consumption requires that it provides the same marginal utility as private consumption. Equation (9) determines the first-best labor supply by equating the marginal rate of substitution of leisure for consumption to the marginal rate of transformation, which is equal to the technology shock.

3.2 Second-best problem

We follow the primal approach of Lucas and Stokey (1983) and express prices and tax rates in terms of allocations by using (6) and (7). As usual, the Ramsey problem can be stated as follows:

Definition 1. *The Ramsey problem is to choose at $t = 0$ $c_t, g_t \geq 0, h_t \in [0,1]$ in order to maximize (2) subject to the implementability constraint*

$$\sum_{t=0}^{\infty} \beta^t \sum_{s^t} \pi_t(s^t)\Omega(c_t(s^t), h_t(s^t), g_t(s^t)) = U_c(c_0, 1-h_0, g_0)b_0. \tag{10}$$

and the resource constraint (1), where (s_0, b_0) given. Ω is defined as $\Omega(c,h,g) \equiv U_c(c,1-h,g)c - U_l(c,1-h,g)h$ and stands for consumption net of after-tax labor income, or, equivalently, primary surplus, in marginal utility of consumption units.

Let Φ denote the multiplier on the unique implementability constraint. We call Φ the excess burden of taxation throughout the paper. Define also

$$\chi \equiv \frac{U_g}{U_c} - 1. \tag{11}$$

We call χ the *public wedge*, since it captures the deviation of the marginal rate of substitution of government consumption for private consumption from its first-best value. We summarize the basic results of the full confidence problem in terms of two propositions.

Proposition 1. *The optimal allocation (c, h, g) is history-independent. Thus, the optimal public wedge and labor tax are history-independent.*

Proof. See the Appendix. □

Proposition 2. *1. The optimal public wedge for $t \geq 1$ is*

$$\chi = \frac{\Phi(1 - \epsilon_{cc} - \epsilon_{ch} - \epsilon_{gc} - \epsilon_{gh})}{1 + \Phi(\epsilon_{gc} + \epsilon_{gh})},$$

where $\epsilon_{cc} \equiv -U_{cc}c/U_c, \epsilon_{ch} \equiv U_{cl}h/U_c$, the own and cross elasticity (with respect to labor) of the marginal utility of private consumption, and $\epsilon_{gc} \equiv U_{gc}c/U_g, \epsilon_{gh} \equiv -U_{gl}h/U_g$ the cross elasticities of the marginal utility of government consumption with respect to private consumption and labor.

2. The optimal labor tax for $t \geq 1$ is

$$\tau = \frac{\Phi(\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc})}{1 + \Phi(1 + \epsilon_{hh} + \epsilon_{hc})}$$

where $\epsilon_{hh} \equiv -U_{ll}h/U_l, \epsilon_{hc} \equiv U_{cl}c/U_l$, the own and cross elasticity (with respect to private consumption) of the marginal disutility of labor.

3. The denominators in all expressions are positive, so the sign of the public wedge and the labor tax depends on the sign of the numerators.

Proof. See the Appendix. □

The history independence of proposition 1 refers to the fact that optimal allocations, and therefore policies, are functions only of the current shock s and the *constant* value of the excess burden of taxation Φ. For example, consumption varies only across shocks, $c_t = c(s_t, \Phi)$. Proposition 1 extends the basic result of Lucas and Stokey (1983) to environments with endogenous government consumption.

Proposition 2 expresses the optimal χ and τ as functions of *elasticities* and the excess burden of taxation Φ.[9] Elasticities of marginal utilities show up in the determination of the wedges because

[9]These formulas are in the spirit of the static analysis with exogenous government expenditures of Atkinson and Stiglitz (1972).

they capture how the surplus in marginal utility units Ω –the main ingredient in the calculation of the present discounted values of future surpluses– is affected by the choices of c, h and g. The proposition implies that when elasticities are constant across states and dates, the public wedge and the labor tax become constant since they depend only on the constant excess burden of taxation. In the next section we will consider a utility function that delivers these results.

3.3 Parametric example

Consider the period utility function

$$U = \frac{u^{1-\rho} - 1}{1-\rho} + v(l), \tag{12}$$

where u stands for a composite good of private and government consumption and $v(l)$ for the subutility of leisure. Assume a constant elasticity of substitution (CES) aggregator u

$$u = [(1-\alpha)c^{1-\psi} + \alpha g^{1-\psi}]^{\frac{1}{1-\psi}}, \alpha \in (0,1).$$

We derive results for the public wedge and the share of government consumption in output that hold independently of the functional form of $v(l)$. The homothetic specification in private and government consumption allows us to perform our analysis in terms of ratios. Furthermore, the specification separates between the intertemporal elasticity of substitution (IES), which is controlled by $1/\rho$, and the intratemporal elasticity of substitution between private and government consumption, which is controlled by $1/\psi$. Separating these two attitudes is key for our later analysis since we will show that the qualitative and quantitative properties of the optimal plan under ambiguity depend on the size of ρ relative to unity (and not on ψ). In contrast, the size of the parameter ψ will determine the distortions at the government consumption margin, as we will soon see. We call c and g *substitutes* when $\psi < 1$ and *complements* when $\psi > 1$.[10]

For the utility function in hand the elasticity of the marginal utility of private consumption is a weighted average of ρ and ψ, $\epsilon_{cc} = \lambda_c \rho + (1-\lambda_c)\psi$, and the cross elasticity of the marginal utility of government consumption with respect to private consumption is $\epsilon_{gc} = (\psi - \rho)\lambda_c$, with weight $\lambda_c \equiv (1-\alpha)(\frac{c}{u})^{1-\psi} \in (0,1)$.[11] Therefore, $\epsilon_{cc} + \epsilon_{gc} = \psi$, so the public wedge in proposition 2 becomes

[10] This utility function has been used extensively in macroeconomic setups where government consumption provides utility. Klein et al. (2008) and Bachmann and Bai (2013) use the case of $\rho = \psi = 1$. Empirical public finance studies have also used this specification in order to estimate the degree of substitutability between private and government consumption. See for example Ni (1995).

[11] Use the CES aggregator u to get $1 = \lambda_c + \lambda_g$, with $\lambda_g \equiv \alpha \left(\frac{g}{u}\right)^{1-\psi}$. The weight λ_c simplifies to $1-\alpha$ for the Cobb-Douglas case of $\psi = 1$.

$$\chi = \frac{\Phi(1-\psi)}{1+\Phi(\psi-\rho)\lambda_c}. \tag{13}$$

As stated in proposition 2, the sign of χ is determined by the numerator of (13). Let $\Lambda \equiv g/y$ denote the share of government consumption in output. Its first-best value (corresponding to a zero public wedge) is $\Lambda^{FB} \equiv \frac{\alpha^{1/\psi}}{\alpha^{1/\psi}+(1-\alpha)^{1/\psi}}$. When c and g are *substitutes* ($\psi < 1$), (13) implies that there is a positive public wedge ($\chi > 0$). Thus, the marginal utility of government consumption is higher that the marginal utility of private consumption, which, by concavity of the utility function, implies that the government share is small relative to the first-best, $\Lambda < \Lambda^{FB}$. The opposite happens in the case of *complements* ($\psi > 1$): the public wedge is negative ($\chi < 0$), implying a large share relative to the first-best, $\Lambda > \Lambda^{FB}$. Thus, presumptions that at the second-best the optimal Λ has to be small relative to the first-best because government consumption has to be financed by distortionary taxation are not valid. For the knife-edge Cobb-Douglas case of $\psi = 1$, the planner does *not* distort the government consumption margin and sets a zero public wedge, leading to the first-best government share, $\Lambda = \Lambda^{FB} = \alpha$ (*levels* of g are of course different).

The following proposition summarizes properties of the optimal government share and taxes for $t \geq 1$.

Proposition 3. *1. Assume the homothetic specification in (12). Then,*

(a) The share of government consumption in output is function only of Φ and not of the shocks s, $\Lambda_t = \Lambda(\Phi)$. Thus, Λ is constant across shocks.

(b) For $\psi \geq \rho$ we have $\operatorname{sign} \Lambda'(\Phi) = \operatorname{sign}(\psi - 1)$. More generally, $\operatorname{sign} \Lambda'(0) = \operatorname{sign}(\psi - 1)$.

2. Assume furthermore constant Frisch elasticity, $v(l) = -a_h \frac{(1-l)^{1+\phi_h}}{1+\phi_h} = -a_h \frac{h^{1+\phi_h}}{1+\phi_h}$. Then,

(a) The tax rate is function only of Φ, $\tau_t = \tau(\Phi)$, and therefore is constant across shocks. Thus, the surplus-to-output ratio, $\tau(\Phi) - \Lambda(\Phi)$, is acyclical.

(b) For $\psi = 1$ or $\psi = \rho$ we have $\tau'(\Phi) > 0$. More generally, $\tau'(0) > 0$.

3. ("Optimality of balanced budgets"). Let the utility function be as in (12) with constant Frisch elasticity. If initial debt is zero, then a balanced budget is optimal for every period. The balanced budget τ and Λ do not depend on the stochastic properties of the shocks but only on preference parameters. If initial debt is positive, then surpluses are optimal for each $t \geq 1$, as long as the initial surplus does not cover the initial level of debt.

Proof. See the Appendix. □

Discussion. With the homothetic utility specification, the history-independence of the share Λ specializes to constancy across shocks. If we further assume a constant Frisch elasticity, a comprehensive *perfect* smoothing result emerges: *both* the government share *and* taxes have to be constant. The dynamics of the optimal plan are pretty simple. If the government starts with zero debt, it runs a balanced budget forever. With positive initial debt, the government runs a deficit at $t = 0$ and then a constant surplus-to-output ratio for each $t \geq 1$. There are neither any positive nor any negative trends in public debt.

The proposition shows also formally that the excess burden of taxation should be interpreted as an indicator of *distortions*. This is obvious for the labor supply margin since labor taxes increase as a function of Φ. Regarding government consumption, increases in Φ reduce the share Λ in the case of substitutes ($\psi < 1$), and increase it in the case of complements ($\psi > 1$). Thus, in both cases, the deviation of the share of government consumption from its first-best value becomes larger. We are particularly interested in the excess burden of taxation because it is the key determinant of the dynamics in an environment with doubts about the probability model of technology shocks.

4 Doubts about the probability model

4.1 Preferences

Until now, we have analyzed an economy where agents have full confidence in the probability measure π. Consider now a situation where the household considers π (which we will call from now on the *reference* measure) a good approximation of the true probability measure but entertains fears that π may be misspecified. In order to deal with the possibility of misspecification, the household considers a *set* of alternative probability measures that are close to π in terms of relative entropy. We are making the assumption that these measures are absolutely continuous with respect to π for finite time intervals and express them as a change of measure. More specifically, the non-negative random variable m_{t+1} denotes a change of the conditional measure $\pi_{t+1}(s_{t+1}|s^t)$. In order to be a proper change of measure it has to integrate to unity, $E_t m_{t+1} = 1$. The unconditional change of measure is defined as $M_t \equiv \prod_{i=1}^{t} m_i$, $M_0 \equiv 1$, and is a martingale with respect to π.

We use the multiplier preferences of Hansen and Sargent (2001) in order to capture this *ambiguity* and the household's *aversion* towards it,[12]

$$V_t = U(c_t, 1 - h_t, g_t) + \beta \min_{m_{t+1} \geq 0, E_t m_{t+1} = 1} [E_t m_{t+1} V_{t+1} + \theta E_t m_{t+1} \ln m_{t+1}], \tag{14}$$

where $\theta > 0$. The parameter θ penalizes probability models that are far from the reference model in terms of relative entropy. Full confidence in the model, and therefore (subjective) expected

[12] See Strzalecki (2011) for a decision-theoretic foundation of the multiplier preferences.

utility is captured by $\theta = \infty$.

4.2 Competitive equilibrium under ambiguity

The cautious household forms worst-case scenarios subject to the entropy penalty. Solving the minimization operation in (14) delivers the worst-case conditional change of measure

$$m_{t+1}(s^{t+1}) = \frac{\exp(\sigma V_{t+1}(s^{t+1}))}{\sum_{s_{t+1}} \pi_{t+1}(s^{t+1}|s^t)\exp(\sigma V_{t+1}(s^{t+1}))} \tag{15}$$

where $\sigma \equiv -\theta^{-1} < 0$, with $\sigma = 0$ corresponding to the expected utility case. Expression (15) shows that an ambiguity averse household assigns higher probability than the reference measure on events that bear low continuation utility and smaller probability than the reference measure on events with high continuation utility. It is important to note that the household's pessimistic beliefs are *endogenous*, since they depend on continuation utility. Using the worst-case model (15) in (14) delivers the familiar risk-sensitive recursion of Tallarini (2000),

$$V_t = U(c_t, 1-h_t, g_t) + \frac{\beta}{\sigma} \ln E_t \exp(\sigma V_{t+1}). \tag{16}$$

Besides the preferences aspect, the rest of the competitive equilibrium is standard. The static labor supply condition (6) remains the same. The intertemporal marginal rate of substitution is altered, leading to an optimality condition with respect to Arrow securities that takes the form

$$p_t(s_{t+1}, s^t) = \beta\pi_{t+1}(s_{t+1}|s^t)m_{t+1}(s^{t+1})\frac{U_c(s^{t+1})}{U_c(s^t)}. \tag{17}$$

The expression for the equilibrium price of an Arrow security provides the connection between the household's endogenous pessimistic beliefs and the fiscal instruments of the planner, which is at the heart of the optimal policy problem: future tax policies affect future utilities and therefore, through the household's endogenous beliefs, equilibrium prices. In turn, equilibrium prices determine the desirability of debt and thus, the trade-off between current taxation and new debt issuance.

4.3 Ramsey problem

As in the case of full confidence in the model, the Ramsey planner chooses the competitive equilibrium that maximizes the utility of the representative household.[13] We follow a recursive represen-

[13]The first-best allocation with doubts about the probability model of technology shocks is the same as with full confidence in the model, due to the essentially static nature of the problem. Let V_0 denote the utility index at $t = 0$. The first-best is characterized by $(-\partial V_0/\partial h_t(s^t))/(\partial V_0/\partial c_t(s^t)) = s_t$ and $(\partial V_0/\partial g_t(s^t))/(\partial V_0/\partial c_t(s^t)) = 1$. For the

tation of the commitment problem from period one onward as in the recursive utility analysis of Karantounias (2013b). Let $z \equiv U_c b$ denote debt in marginal utility units. Debt in marginal utility units lives in the set $Z(s)$ when the current shock is s. Let $V(z,s)$ denote the value function of the planner. Assume that shocks are Markov with transition density $\pi(s'|s)$. Then V is described by the following Bellman equation:

$$V(z,s) = \max_{c,h,g,z'_{s'}} U(c,1-h,g) + \frac{\beta}{\sigma} \ln \sum_{s'} \pi(s'|s) \exp\big(\sigma V(z'_{s'},s')\big)$$

subject to

$$z = \Omega(c,h,g) + \beta \sum_{s'} \pi(s'|s) \frac{\exp(\sigma V(z'_{s'},s'))}{\sum_{s'} \pi(s'|s) \exp\big(\sigma V(z'_{s'},s')\big)} z'_{s'} \tag{18}$$

$$c + g = sh \tag{19}$$

$$c, g \geq 0, h \in [0,1], z'_{s'} \in Z(s') \tag{20}$$

Let Φ denote the multiplier on the dynamic implementability constraint (18) and let λ denote the multiplier on the resource constraint (19). The first-order necessary conditions are

$$c: \quad U_c + \Phi\Omega_c = \lambda \tag{21}$$

$$h: \quad -U_l + \Phi\Omega_h = -\lambda s \tag{22}$$

$$g: \quad U_g + \Phi\Omega_g = \lambda \tag{23}$$

$$z'_{s'}: \quad V_z(z'_{s'},s')[1 + \sigma\eta'_{s'}\Phi] + \Phi = 0 \tag{24}$$

where $\eta'_{s'} \equiv z'_{s'} - \sum_{s'} \pi(s'|s) m'_{s'} z'_{s'}$. The variable $m'_{s'}$ stands for the conditional likelihood ratio, $m'_{s'} = \frac{\exp(\sigma V(z'_{s'},s'))}{\sum_{s'} \pi(s'|s) \exp(\sigma V(z'_{s'},s'))}$. $\Omega_i, i = c,h,g$ stands for the respective partial derivative of the surplus in marginal utility units Ω.

We call the variable $\eta'_{s'}$ the *relative* debt position in marginal utility units, since it denotes the size of $z'_{s'}$ with respect to the "average" debt position. The relative debt position can be positive ($\eta'_{s'} > 0$) or negative ($\eta'_{s'} < 0$). Furthermore, it is on average *zero* under the worst-case model, i.e. $\sum_{s'} \pi(s'|s) m'_{s'} \eta_{s'} = \sum_{s'} \pi(s'|s) m'_{s'} z'_{s'} - \sum_{s'} \pi(s'|s) m'_{s'} (\sum_{s'} \pi(s'|s) m'_{s'} z'_{s'}) = 0$, since $\sum_{s'} \pi(s'|s) m'_{s'} = 1$.

multiplier preferences we have $\partial V_0/\partial h_t(s^t) = -\beta^t \pi_t M_t U_l(s^t)$, $\partial V_0/\partial c_t(s^t) = \beta^t \pi_t M_t U_c(s^t)$ and $\partial V_0/\partial g_t(s^t) = \beta^t \pi_t M_t U_g(s^t)$, which lead to (8) and (9).

4.4 Initial remarks

The important element that doubts about the model contribute is an excess burden of taxation that is *not* constant anymore. In particular, use the envelope condition $V_z(z,s) = -\Phi$ and rewrite (24) in sequence notation as

$$\frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} + \sigma\eta_{t+1}, t \geq 0 \tag{25}$$

where $\eta_{t+1} = z_{t+1} - E_t m_{t+1} z_{t+1}$. The law of motion (25) will be analyzed in detail in the next section. It obviously implies that with full confidence in the model ($\sigma = 0$), we have $\Phi_t = \Phi \forall t, s^t$.

Following the same steps as in the proofs of propositions 1 and 2, we can write the optimal allocation (c, h, g) (and therefore the optimal policy instruments τ and Λ) as functions of the current shock s_t and the time-varying excess burden, Φ_t. These functions are exactly the same functions of (s, Φ) as in the case without doubts about the model. As a result we have:

Proposition 4. *(Optimal wedges with doubts about the model) The optimal public wedge and the optimal tax rate are as in proposition 2, with an excess burden of taxation that follows now the law of motion (25). Thus, all formulas for our parametric example in proposition 3 go through by replacing Φ with Φ_t. The optimal τ and Λ will not be constant anymore, but they will reflect the variation in Φ_t.*

5 Fiscal policies over states and dates

The goal of the rest of the paper is to understand the dynamics of Φ_t, which determine the dynamics of optimal taxes and government consumption.

5.1 Excess burden of taxation and debt in marginal utility units

With doubts about the model, the excess burden of taxation Φ_t depends on the relative debt position in marginal units η_{t+1}. The law of motion (25) implies that the excess burden increases ($\Phi_{t+1} > \Phi_t$) when there is a positive relative debt position $\eta_{t+1} > 0$, i.e. when debt in marginal utility units z_{t+1} is larger than the average position $E_t m_{t+1} z_{t+1}$, and it decreases ($\Phi_{t+1} < \Phi_t$), when there is a negative relative position, $\eta_{t+1} < 0$, so when z_{t+1} is smaller than the average position.

These changes in the excess burden of taxation take place because the endogenous household's beliefs are the source of a novel *price* effect that the policymaker is manipulating in order to make debt less costly. To see that, consider an increase in the state-contingent position $z'_{s'}$ at s'. More debt decreases utility (since debt has to be repaid with distortionary taxation) and, as a result, it increases the probability that the pessimistic household assigns to this state of the world, according to (15). Thus, the respective Arrow security becomes more expensive, as can been seen from (17). This increase in price is beneficial to the planner if he takes a positive relative debt position, since the price at which he *sells* debt increases (and therefore the state-contingent return on debt falls) and harmful in the opposite case. Therefore, instead of keeping the excess burden constant over states and dates, the planner increases the excess burden of taxation at states of the world next period against which it is cheaper to issue debt, and decreases distortions at states of the world for which debt is relatively expensive.[14]

An equivalent, perhaps more intuitive, interpretation of the law of motion of the excess burden is available if we thought in terms of the policy instrument of the planner, say tax rates τ_{t+1}. High tax rates decrease the utility of the household and increase equilibrium prices through the pessimistic beliefs. By increasing future taxes against states of the world for which z_{t+1} is high and reducing future taxes against states of the world where z_{t+1} is low, the overall value of the portfolio of new state-contingent claims increases, an outcome which relaxes the current government budget constraint and increases therefore welfare.

Finally, note from (25) that Φ_t remains constant if the relative debt positions are zero for all dates and states, $\eta_{t+1} = 0, t \geq 0$. That is, the price manipulation mechanism through the endogenous beliefs is relevant only if state-contingent debt in marginal utility units *does* vary across shocks or if it is actually necessary to issue debt. Otherwise, the effect of model uncertainty on policy is *muted*, and the full confidence fiscal plan is followed:

Proposition 5. *("Muting the effect of doubts on policy") Let $\Omega^\star(s,\Phi)$ denote the optimal surplus in marginal utility units as a function of (s,Φ).*

1. *Assume that $\Omega^\star(s,\Phi) = \Omega^\star(s',\Phi), \forall \Phi, \forall s \neq s'$. Then $\Phi_t = \Phi^{no\ doubts}$, where $\Phi^{no\ doubts}$ is the excess burden of taxation of the economy with full confidence in the model.*

2. *Assume that $b_0 = 0$ and that there exists a $\bar{\Phi}$ such that $\Omega^\star(s,\bar{\Phi}) = \Omega^\star(s',\bar{\Phi}) = 0, \forall s \neq s'$. Then $\Phi_t = \bar{\Phi}$ and the planner runs the same balanced budget as in an economy without doubts.*

In both cases, the allocation (c,h,g) and policies (τ,Λ) are the same as in an economy without doubts. Only equilibrium asset prices are different.

Proof. See the Appendix. □

The set of period utility functions that generate the above results is not empty:

Corollary. *If 1) the utility function is as in (12) with $\rho = 1$ and we have any subutility of leisure $v(l)$ or if 2) initial debt is zero and the utility function is as in (12) with constant Frisch elasticity, then doubts about the model leave the second-best allocation and policies unaltered.*

[14] This mechanism has been previously partially uncovered in Karantounias (2013a), where it was counteracted by a paternalistic incentive of the planner, and is present – for different reasons– in environments with preference for early resolution of uncertainty, as in Karantounias (2013b). See the related literature section in the Introduction.

Proof. See the Appendix. □

5.2 Managing expectations and the prominent role of the IES

The previous section shows that the planner manages the pessimistic expectations of the household with an ultimate goal of making debt cheaper, improving the trade-off between taxing today versus issuing state-contingent debt, i.e. versus taxing in the future. The planner always taxes more against states of the world for which the "value" of debt, i.e. debt in *marginal utility* units, z_{t+1}, is high. Such a taxation scheme *amplifies* the market value of government debt since a high tax increases equilibrium prices and increases therefore the *revenue* from new debt issuance.

But why is debt in marginal utility units the relevant object for the increase of excess burden and not just debt? The reason is squarely based on the logic of the intertemporal budget constraint of the government. The present discounted value of future surpluses entails both an adjustment for model uncertainty, through the pessimistic expectations, and an adjustment for risk, through marginal utility. The planner uses the new tool, the pessimistic expectations, to increase essentially the weight on high "values" of debt in order to make the present value of surpluses higher, and relax therefore the fiscal constraint.

Thus, the response of the "value" of debt to the cycle is our ultimate object of interest. We will provide a sharp characterization of the value of debt in terms of the intertemporal elasticity of substitution, which determines the elasticity of marginal utility and therefore, the reaction (partially) of interest rates to technology shocks. We will show that when the IES is smaller than unity (so $\rho > 1$), then values of debt (z_{t+1}) are high when technology shocks are low, leading to high taxes in recessions, or otherwise "austerity" policies. In contrast, when the IES is large and therefore equilibrium interest rates not so responsive, then values of debt are high when technology shocks are high, leading to high taxes in good times, or "anti-austerity" policies.

5.2.1 The role of the IES in a two-period economy

To see clearly the mechanism, proceed first to a two-period version of our economy. Debt in marginal utility units z simplifies to surplus in marginal utility units, Ω, a fact which allows a simple characterization for *small* doubts about the model.

Proposition 6. *Assume that shocks take two values, $s_L < s_H$ and let $i = L, H$ denote the state of the world at $t = 1$ with an excess burden of taxation*

$$\frac{1}{\Phi_i} = \frac{1}{\Phi_0} + \sigma\eta_i \quad \textit{where} \quad \eta_i = \Omega_i - \sum_i \pi_i m_i \Omega_i, i = L, H.$$

Let $\Omega_i^{\sigma=0}, i = L, H$ denote the surplus in marginal utility units that pertains to the full confidence analysis, $\sigma = 0$. Then:

1. *If $\Omega_H^{\sigma=0} > \Omega_L^{\sigma=0}$, then $\Phi_H > \Phi_0 > \Phi_L$ for small σ. If $\Omega_H^{\sigma=0} < \Omega_L^{\sigma=0}$, then $\Phi_H < \Phi_0 < \Phi_L$ for small σ.*

2. *Let the utility function be as in (12) with constant Frisch elasticity. Then, the optimal surplus in marginal utility units as function of (s, Φ) takes the following form:*

$$\Omega^\star(s, \Phi) = \big(\tau(\Phi) - \Lambda(\Phi)\big) J(\Phi) \cdot \big[y(s, \Phi)\big]^{1-\rho}, \tag{26}$$

where $J(.) > 0$ defined in the Appendix. Assume that we have surpluses at $t = 1$ for $\sigma = 0$, $\tau > \Lambda$. Expression (26) implies that when $\sigma = 0$, the surplus in marginal utility units is countercyclical if $\rho > 1$ ($\Omega_H^{\sigma=0} < \Omega_L^{\sigma=0}$), and procyclical ($\Omega_H^{\sigma=0} > \Omega_L^{\sigma=0}$) if $\rho < 1$. Thus, part 1) implies that for small σ, the excess burden is countercyclical ($\Phi_H < \Phi_L$) if $\rho > 1$ and procyclical ($\Phi_H > \Phi_L$) if $\rho < 1$.

Proof. See the Appendix. □

The proposition allows us to use the behavior of surpluses in the full confidence economy as an indicator of distortions in the economy with doubts. This is neat because the full confidence economy is easily characterized, due to a constant excess burden.

Expression (26) shows that the surplus in marginal utility units is proportional to a multiple of output in the power of $(1 - \rho)$. Keeping Φ constant (which implies a constant policy τ *and* Λ), surpluses are procyclical, since a positive technology shock leads to output expansion. An expansion in output and therefore an expansion in consumption is counteracted though by a *contraction* in marginal utility, which is controlled by ρ, the inverse of IES. If $\rho > 1$, marginal utility is very elastic and therefore the negative marginal utility effect dominates, offsetting the cyclicality of surpluses. As a result, surpluses in *marginal utility* units become *countercyclical*, leading to a countercyclical excess burden. In contrast, when $\rho < 1$ and therefore when the intertemporal elasticity of substitution is high ($IES > 1$), surpluses in marginal utility units, and thus, the excess burden, remain procyclical.[15]

Therefore, the role of the IES with ambiguity aversion is to indicate which states of the world have the most potential for interest rate manipulation through the channel of the endogenous pessimistic beliefs. Tax rates follow the same pattern as the excess burden of taxation. Hence, when the IES is smaller than unity, "austerity" measures become *optimal*: taxes increase in bad times and decrease in good times, amplifying the cycle. In contrast, a high IES leads to an "anti-austerity" policy: taxes are high in good times and low in bad times, attenuating the cycle.

[15]Note that for $\rho = 1$, Ω stays constant across shocks, as we expect from the corollary of proposition 5. Note also that the response of Ω is controlled only by ρ for constant Φ, despite the fact that U_c depends on both ψ and ρ (recall that $1/\rho$ captures the IES of the composite good u). The reason behind that is the fact that ψ affects the results through the determination of the government share Λ (together with ρ), which enters through the function $J(\Phi)$. For constant Φ, Λ becomes constant, leading to a clean dependence of the cyclicality of Ω on ρ.

Intuitively, the IES matters because it captures the equilibrium interest rate sensitivity to shocks. When the marginal utility is very elastic, $\rho > 1$, and therefore when the household does not substitute a lot intertemporally, shocks to output and consumption induce a big change to the intertemporal rate of substitution and therefore big changes to interest rates. A bad shock tomorrow implies –through the increase in marginal utility– a decrease in the state-contingent return of debt. If this decrease in interest rates is large enough (which is the case when $\rho > 1$), then the planner has an incentive to amplify the reduction through the pessimistic beliefs of the household: he issues more state-contingent debt and taxes more against bad shocks. The opposite happens for a good technology shock tomorrow. The planner taxes less against the good shock and issues less debt, since the state-contingent yield rises too much when $\rho > 1$, and therefore the revenue from debt issuance against a good shock decreases.

5.2.2 Infinite horizon and IES near the balanced budget

In an infinite horizon economy the behavior of debt in marginal utility units becomes more entangled since z_{t+1} depends on both surpluses in marginal utility units and on the pessimistic beliefs. We are able though to fully characterize optimal policy by focusing on the vicinity of the balanced budget.

Consider any period utility function that delivers a tax rate and a government share that are functions only of Φ. This type of utility function exhibits the convenient feature that the level of excess burden of taxation Φ^* that delivers a balanced budget, $\tau(\Phi^*) = \Lambda(\Phi^*)$, is a *fixed point* of the law of motion (25). If this point is ever reached, i.e. if government debt becomes zero, then it is optimal to run a balanced budget forever, with doubts about the model affecting only asset prices, as shown in proposition 5. Our strategy here is to treat Φ_t as a state variable with law of motion (25) and proceed with an approximation of the equilibrium around Φ^*.

Let the shocks take N values, and let them be enumerated by the index i from the smallest to the largest. Let $\Phi_{j|i}(\Phi)$ denote the excess burden of taxation next period at the realization of the technology shock j when the current shock and excess burden are i and Φ respectively. The approximate law of motion of the excess burden takes the form

$$\Phi_{j|i}(\Phi) \simeq \Phi^* + \Phi'_{j|i}(\Phi^*)(\Phi - \Phi^*), \quad i, j = 1, ..., N. \tag{27}$$

The object of interest is $\Phi'_{j|i}(\Phi^*)$, which stands for the derivative of the excess burden, when we have the transition from i to j, evaluated at the balanced budget.[16] Let $m^*_{j|i}$ denote the conditional likelihood ratio from i to j evaluated at the balanced budget. Let $\Pi \equiv [\pi(j|i)]$ denote the transition matrix of the Markov process and let $\mathbf{M} \equiv [m^*_{j|i}]$, $\boldsymbol{\Phi} \equiv [\Phi'_{j|i}(\Phi^*)]$ denote the matrices

[16] Note that in contrast to typical approximations around the deterministic steady state, we do not turn off uncertainty, a fact which makes the derivatives $\Phi'_{j|i}$ depend on shocks.

that collect $m^*_{j|i}$ and $\Phi'_{j|i}(\Phi^*)$ at the i-th row and j-th column. We have the following properties for the approximate excess burden of taxation.

Proposition 7. *("Properties of the excess burden of taxation near the balanced budget")*

1. *("Monotonicity") Let the current excess burden of taxation be larger than its balanced budget value, $\Phi > \Phi^*$, which is the case when we have positive debt. Then, (27) implies that*

   - *If $\Phi'_{j|i}(\Phi^*) > (<)1 \Rightarrow \Phi_{j|i}(\Phi) > (<)\Phi$.*
   - *If $\Phi'_{k|i}(\Phi^*) > (<)\Phi'_{l|i}(\Phi^*) \Rightarrow \Phi_{k|i}(\Phi) > (<)\Phi_{l|i}(\Phi)\forall k, l, i$.*

2. *("Martingale") We have*

$$\sum_j \pi(j|i) m^*_{j|i} \Phi'_{j|i}(\Phi^*) = 1, \forall i. \tag{28}$$

   Property (28) implies that the excess burden of taxation is a martingale with respect to the worst-case measure, $\sum_j \pi(j|i) m^*_{j|i} \Phi_{j|i}(\Phi) = \Phi, \forall i$. Thus, the matrix $\mathbf{A} \equiv \Pi \circ \mathbf{M} \circ \mathbf{\Phi}$, where $\circ$ denotes element-by-element multiplication, is stochastic.

3. *("Drifts") When $m^*_{j|i}$ decreasing in j (i.e. when the household's worst case model puts less probability mass on high technology shocks), we have:*

   - *If $\Phi'_{j|i}(\Phi^*)$ is decreasing in j, then $\sum_j \pi(j|i)\Phi'_{j|i}(\Phi^*) < 1$. This implies $\sum_j \pi(j|i)\Phi_{j|i}(\Phi) < \Phi$ when $\Phi > \Phi^*$, so there is a negative drift with respect to π when we start with positive debt.*
   - *If $\Phi'_{j|i}(\Phi^*)$ is increasing in j, then $\sum_j \pi(j|i)\Phi'_{j|i}(\Phi^*) > 1$. This implies $\sum_j \pi(j|i)\Phi_{j|i}(\Phi) > \Phi$ when $\Phi > \Phi^*$, so there is a positive drift with respect to π when we start with positive debt.*

Proposition 7 shows that we can characterize the cyclicality and drifts of the excess burden of taxation by considering the entries of the matrix $\mathbf{\Phi}$. If for each row i, $\Phi'_{j|i}(\Phi^*)$ are decreasing in j, then their weighted average according to the reference model π is smaller than unity (since the non-pessimistic reference model assigns smaller probability mass on low technology shocks that bear high excess burden) and we have *both* countercyclicality of distortions *and* a negative drift with respect to π. In contrast, if the entries of each row i are increasing in the column j, we have procyclicality of distortions and a positive drift with respect to π. Note that to derive proposition 7, we only assumed that (τ, Λ) depend solely on Φ. The next proposition considers parametric forms that deliver this assumption and utilizes proposition 7 by connecting the monotonicity of $\Phi'_{j|i}(\Phi^*)$ to the IES.

Proposition 8. *("IES and austerity near the balanced budget")*

1. *Let the utility function be as in (12) with constant Frisch elasticity and assume that $N = 2$. If $\tau'(\Phi^*) > \Lambda'(\Phi^*)$ we have the following:*
   - *(**"Austerity"**) Assume that $\rho > 1$. Then, $\Phi'_{1|i}(\Phi^*) > 1 > \Phi'_{2|i}(\Phi^*), i = 1, 2$. So the excess burden is countercyclical. Furthermore, the excess burden exhibits a negative drift with respect to π when $\Phi > \Phi^*$, if $m^*_{j|i}$ is decreasing in j $\forall i$.*
   - *(**"Anti-austerity"**) Assume that $\rho < 1$. Then, $\Phi'_{1|i}(\Phi^*) < 1 < \Phi'_{2|i}(\Phi^*), i = 1, 2$. So the excess burden is procyclical. Furthermore, the excess burden exhibits a positive drift with respect to π when $\Phi > \Phi^*$, if $m^*_{j|i}$ is decreasing in j $\forall i$.*
2. *Assume the balanced-growth consistent preferences $U = \frac{u^{1-\rho}-1}{1-\rho}$, where $u = c^{\alpha_1} l^{\alpha_2} g^{\alpha_3}, \alpha_i > 0, \sum_i \alpha_i = 1$. Both the results of the two-period economy of proposition 6 and the results of part (1) of the current proposition go through.*

Proof. See the Online Appendix for the derivations behind propositions 7 and 8. □

Discussion. Proposition 8 generalizes the two-period results of proposition 6 to an infinite horizon setup.[17] Furthermore, it shows that our results hold for a broader set of preferences that satisfy balanced-growth restrictions. The same mechanisms are in play as in the two-period model; the IES controls the sensitivity of z with respect to shocks in the expected way. High elasticity of marginal utility offsets the procyclicality of debt, making debt in marginal utility units, and therefore the excess burden, countercyclical.

The new element that arises in infinite horizon involves *drifts* in the excess burden of taxation, which are non-existent in full confidence economies. This is not an arbitrary consequence of the balanced budget approximation in propositions 7 and 8, but a general feature of the policy problem. As in recursive utility environments like Karantounias (2013b), the inverse of Φ_t is a martingale with respect to the worst-case measure in the non-linear economy, due to the fact that the average relative debt position is zero, $E_t m_{t+1} \eta_{t+1} = 0$. Therefore, the excess burden of taxation is a submartingale with respect to the worst-case measure, $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$. So, distortions increase on average over time with respect to the pessimistic beliefs. The drift with respect to π depends on the conditional covariance of the household's worst-case beliefs m_{t+1} with Φ_{t+1}, since $E_t \Phi_{t+1} \geq \Phi_t - Cov_t(m_{t+1}, \Phi_{t+1})$. The covariance is positive if $\rho > 1$, opening the possibility of a negative drift with respect to π, and negative if $\rho < 1$, maintaining the positive drift.[18]

[17] The restriction that the slope of the tax schedule at the balanced budget is larger than the slope of the share of government expenditures ($\tau'(\Phi^*) > \Lambda'(\Phi^*)$) is similar to the assumption in proposition 6 that $\tau > \Lambda$ and is typically satisfied for our parametric examples.

[18] Recall that for $\rho > 1$ the excess burden increases in bad times, leading to a positive covariance, since bad times are weighed more by the pessimistic household. The opposite happens when $\rho < 1$.

To summarize, two different cases emerge:

- *Front-loading* of distortions when $IES < 1$ ($\rho > 1$): distortions are countercyclical. The planner *both* increases taxes in bad times, *and* decreases on average taxes over time if uncertainty is actually generated by π, so a front-loaded *fiscal consolidation* becomes optimal. We expect decumulation of government debt, until it becomes zero and the primary budget is balanced.

- *Back-loading* of distortions when $IES > 1$ ($\rho < 1$): distortions are procyclical and there is a positive drift in Φ_t if π drives uncertainty. This back-loading of distortions implies that the tax rate and government debt increase on average over time.

These two cases will be valid also outside the vicinity of the balanced budget, as we numerically show in the next section.

6 Numerical simulations

Besides being helpful for deriving theoretical results, the approximate law of motion (27) can be used also for the numerical solution of the problem, as long as we constrain ourselves to the vicinity of the balanced budget.[19] We are interested here in the case where initial debt is large, so we will resort to a global solution method.

The global solution of the problem is non-trivial due to the presence of the value functions in the constraints which hinder the contraction property. We provide details about our solution method in the Appendix.

6.1 Calibration

We use a standard calibration for our parametric example (12) with a constant Frisch elasticity. We set $(\beta, \phi_h) = (0.96, 1)$ to get an annual frequency and a unitary Frisch elasticity. Let the logarithm of technology shocks $a_t \equiv \ln s_t$ follow an AR(1) process, $a_t = \rho_a a_{t-1} + \epsilon_t$, with $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. We set the persistence parameter to $\rho_a = 0.95^4 = 0.8145$ and $\sigma_\epsilon = 0.0174$. These values imply a 3% unconditional standard deviation of the technology shock, $\sigma_a = 0.03$. We take the stance that this autoregressive process, which is the reference model π that the household doubts, is also the *true* data-generating process, i.e. the household's fears of model misspecification are unfounded. We approximate the AR(1) process with two points using the Rouwenhorst method of Kopecky and Suen (2010) and get $(s_L, s_H) = (0.9704, 1.0305)$, and $\pi(i|i) = 0.9073, i = L, H$.

The crucial parameter for the allocation of distortions with doubts about the model is ρ. We set $\rho = 2$ for our baseline calibration and consider also the case of a high IES with $\rho = 0.5$. For

[19] The Online Appendix details an algorithm for finding the matrix $\mathbf{\Phi}$. Results using the expansion are available upon request.

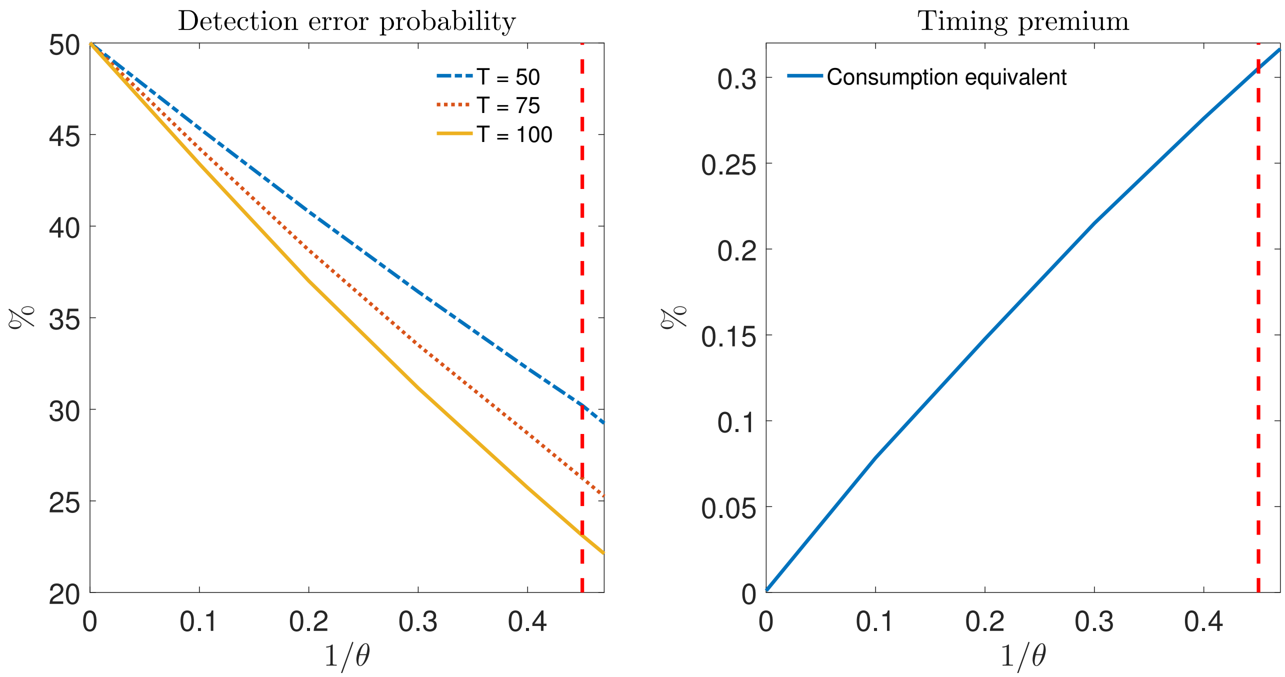

Figure 1: The left graph depicts p as function of $1/\theta$ for $(\rho, \psi) = (2, 1)$. The longer the length of the sample, the smaller the detection error probability for a given level of σ. For each θ 100,000 sample paths were generated according to the reference and the worst-case model. The vertical dotted line at $\sigma = -0.45$ corresponds to $p = 30.21\%$ for $T = 50$, $p = 26.24\%$ for $T = 75$ and $p = 23.11\%$ for $T = 100$. The right graph depicts the timing premium as function of $1/\theta$ for $(\rho, \psi) = (2, 1)$. For each θ the premium was computed using a fixed time horizon of $T = 1,000$ years and 1,000 simulations. The vertical dotted line at $\sigma = -0.45$ corresponds to a timing premium of 0.3025%.

our baseline analysis we set $\psi = 1$, which implies a zero public wedge and constant government share that is independent of Φ. We explore later the implications of model uncertainty on Λ. We set $(\alpha, a_h) = (0.2, 25.77)$ so that Λ is 20% and the household works 40% at the first-best when shocks are at their average value of unity. The initial shock is $s_0 = s_L$, and initial debt is $b_0 = 0.2$, which corresponds to 50% of first-best output.

Detection error probabilities. We discipline the choice of $\sigma \equiv -1/\theta$, the parameter that captures the decision maker's doubts about π, by using the detection error probabilities methodology of Anderson et al. (2003).[20] The detection error probability stands for the probability of rejecting a particular model with a likelihood ratio test, when this model is actually the true data-generating process. Probability models that are "close" to each other imply a high probability of a detection error. The further apart two models are, so the higher σ in absolute value, the easier it is to statistically distinguish them, and the lower the detection error probability.

Let model A and model B stand for the reference model π and the worst-case model respectively, and remember that M_t stands for the unconditional likelihood ratio of the worst-case model to the reference model. The detection error probabilities for the two models for data of length T are

[20] See also Hansen and Sargent (2008) and Barillas et al. (2009) for further examples.

$$
\begin{aligned}
p_A &= Prob(\text{reject A}|\text{data generated by A}) = Prob(M_T > 1|\text{data generated by A}) \\
p_B &= Prob(\text{reject B}|\text{data generated by B}) = Prob(M_T < 1|\text{data generated by B}).
\end{aligned}
$$

If we think that the two models are a priori equiprobable, then the detection error probability is $p = 0.5 \cdot p_A + 0.5 \cdot p_B$. The left graph in figure 1 plots this probability as function of $1/\theta$ for the baseline scenario of $\rho = 2$. Note that when θ is very high, i.e. when there are small doubts about the model, the two models are essentially the same and the detection error probability becomes close to 50%. The graph plots p for sample paths that are 50, 75 or 100 periods (years) long. We set $\sigma = -0.45$ that corresponds to a detection error probability of 30% when $T = 50$, or 26% and 23% for sample paths of 75 or 100 years length respectively. For the case of high IES, $\rho = 0.5$, we re-calibrate σ to -1.7. This value corresponds to a detection error probability of 41% when $T = 50$.[21] Overall, our choices of σ do not imply large doubts about the model; Hansen and Sargent (2008) regard a detection error probability as low as 10% as justifiable.

Timing premium. The detection error probability exercise treats seriously the notion of model uncertainty for the calibration of the parameter σ. The equivalence of the multiplier preferences (14) with the risk-sensitive recursion (16) allows us to explore a different avenue and associate σ to the *timing premium* of Epstein et al. (2014).[22]

Epstein et al. (2014) define as the timing premium the fraction of the consumption stream that the decision maker would be willing to give up in order for all risk to be resolved at $t = 1$. This is a thought experiment that puts in perspective the strength of preference for early resolution of uncertainty, as captured by preference parameters and the specification of the exogenous stochastic processes. We perform the same exercise, modified appropriately for an economy with production and optimal policy.[23] The right panel of figure 1 plots the timing premium as function of $1/\theta$. The timing premium is zero when $\sigma = 0$. The larger $1/\theta$ – which in a recursive utility world would translate to a larger aversion to future consumption risks– the larger the timing premium. For $\sigma = -0.45$ the timing premium is 0.3%, i.e. the household would give up up to 0.3% of its consumption stream, in order to live in a world where all uncertainty is resolved at $t = 1$. The magnitude of timing premia is small because there is no growth risk in our economy.

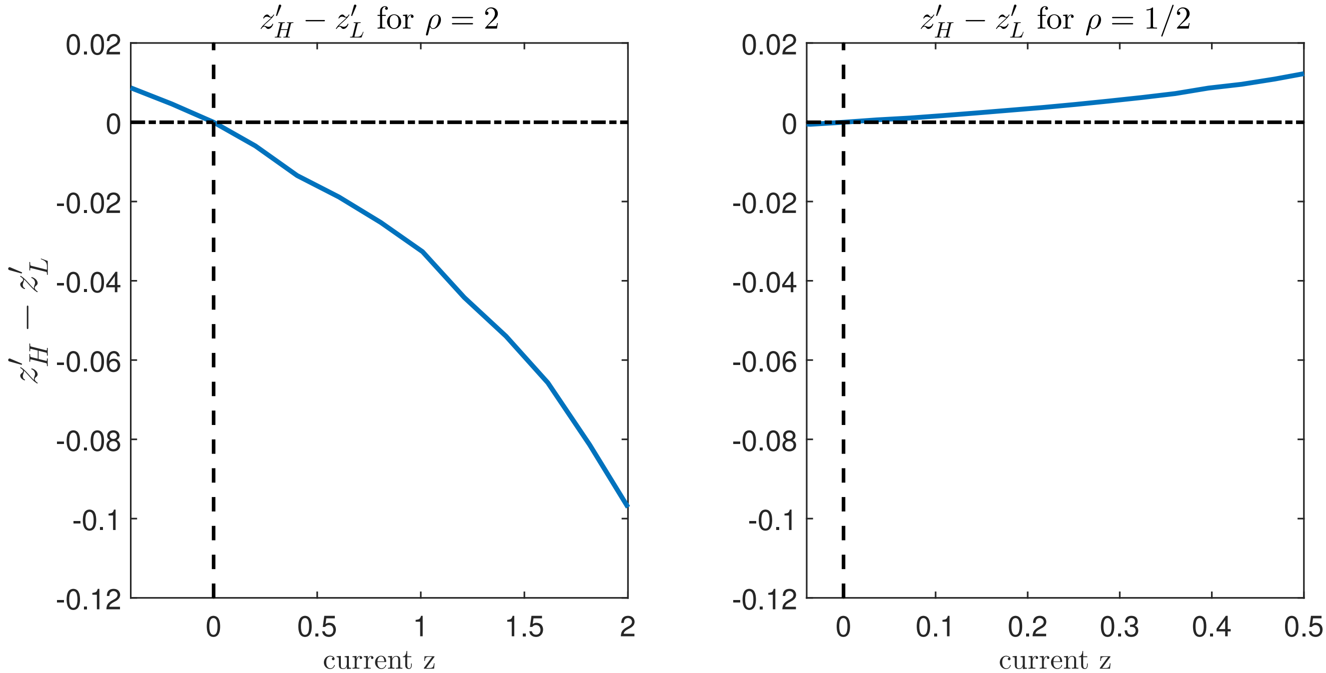

Figure 2: The left graph considers the case of an IES smaller than unity and depicts the difference between the policy functions for debt contingent on a high shock next period minus debt contingent on a low shock next period. The current shock is low, $s = s_L$. The right graph performs the same exercise for an IES larger than unity. Since the government issues debt for $t \geq 1$, the relevant parts of the policy functions are for $z > 0$.

Table 1: Correlations of the excess burden of taxation and the tax rate.

	$\rho = 2$	$\rho = 0.5$
Correlation of $\Delta\Phi$ with s	-0.5138	0.5934
Correlation of $\Delta\Phi$ with y	-0.5054	0.5941
Autocorrelation of Φ	0.9897	0.9792
Correlation of $\Delta\tau$ with s	-0.5145	0.6106
Correlation of $\Delta\tau$ with y	-0.5062	0.6113
Autocorrelation of τ	0.9897	0.9792

The table depicts mean statistics for the cases of low and high IES. We simulated $10,000$ paths and used the first 200 periods of each sample path for the calculation of the respective statistic.

6.2 Policy functions and correlations

Figure 2 plots the policy functions for state-contingent debt in marginal utility units next period for the case of a low and high IES. As expected from the analysis in the previous section, the value

[21]The larger σ is in absolute value, the stronger the non-convexities of the optimal policy problem. Strong non-convexities create convergence problems to our solution algorithm. This is why we refrained from trying to reach a detection error probability of 30% as in the low IES case.

[22]See Strzalecki (2013) for the analysis of ambiguity aversion and the temporal resolution of uncertainty.

[23]The details are provided in the Appendix.

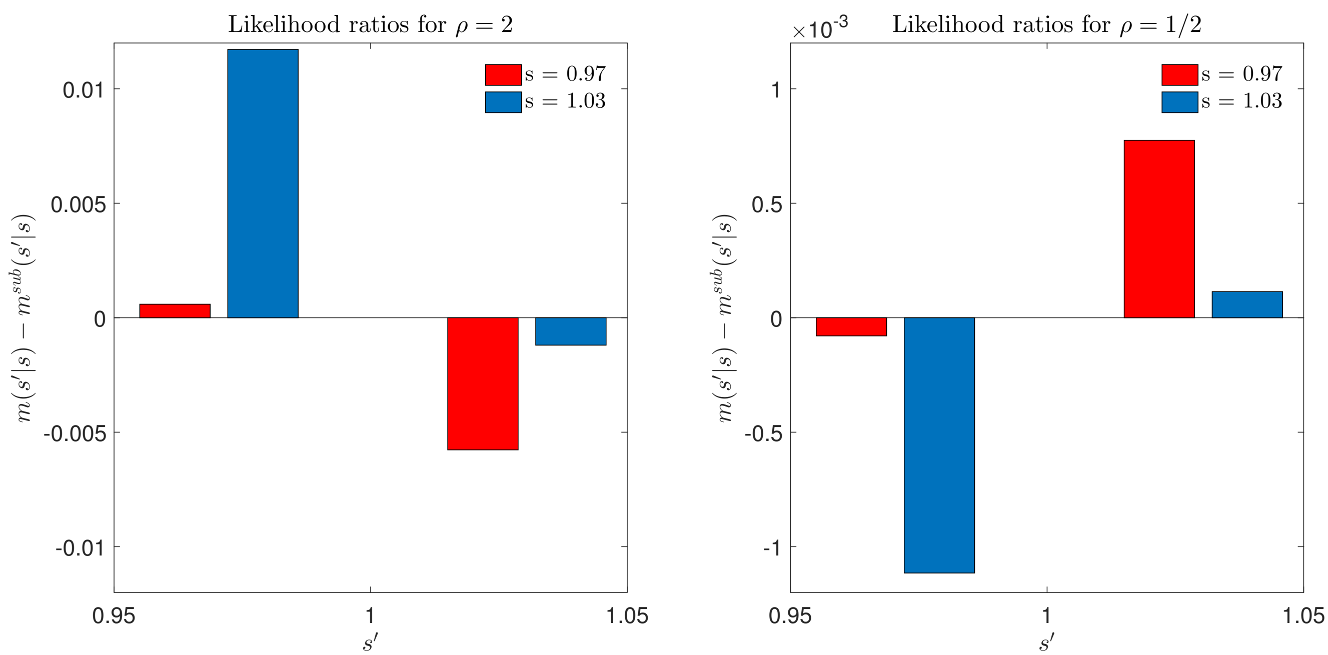

Figure 3: Amplifying pessimistic expectations for IES smaller than unity (left graph) and mitigating pessimistic expectations for IES larger than unity (right graph). Each plot depicts the difference in conditional likelihood ratios, $m(s'|s)$ under optimal policy minus $m^{sub}(s'|s)$ under passive policy.

of debt is countercyclical ($z'_H < z'_L$) when the IES is low and procyclical ($z'_H > z'_L$) when the IES is high. Thus, the results of proposition 8 near the balanced budget extend to the global solution. The excess burden is countercyclical for the low IES case ("the austerity" case) and procyclical for the high IES case. Taxes are a monotonic function of the excess burden of taxation, and, therefore, exhibit the same behavior. Table 1 provides estimates of linear correlation coefficients of the change in the excess burden of taxation and the tax rate with technology shocks and output.

Amplifying versus mitigating pessimistic expectations. The incentives to manage expectations are always associated with the respective benefits of manipulating debt values, which depend on the IES. Figure 3 contrasts the *optimal* conditional likelihood ratio m_{t+1} with the likelihood ratio that would emerge if the planner did not recognize the effects of the endogenous beliefs on asset prices and followed a "passive" policy of a constant excess burden of taxation. The optimal policy prescribes to tax more in bad (good) times when the IES is lower (larger) than unity. Thus, relative to the passive policy, the planner is either amplifying the pessimistic expectations by decreasing utility more in bad times through a higher tax ($IES < 1$), or he is mitigating the pessimism of the household by reducing taxes in bad times ($IES > 1$). These small differences between the passive and the optimal pessimistic beliefs actually result in great differences in the long-run dynamics of optimal policy, as we show in the next section.

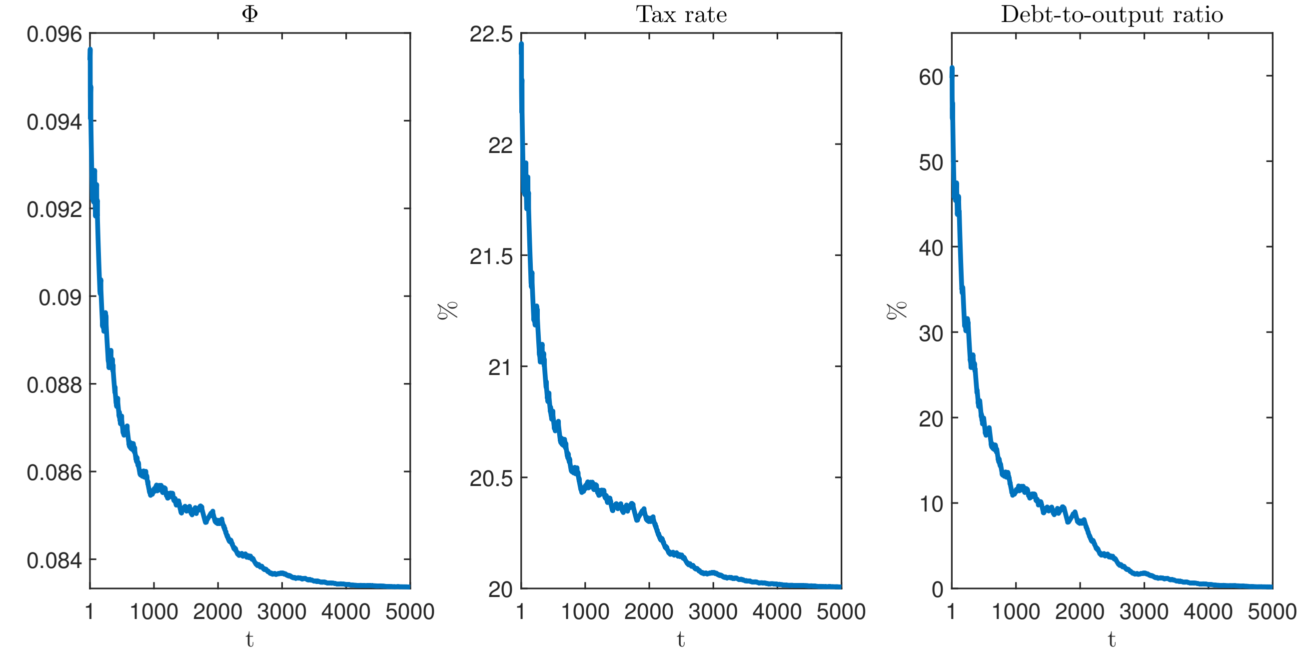

Figure 4: Typical sample path for $\rho = 2$. It displays long-run convergence to the balanced primary budget with zero public debt. The balanced budget tax rate is 20%. The government runs at $t = 0$ a deficit that is 6.15% of output.

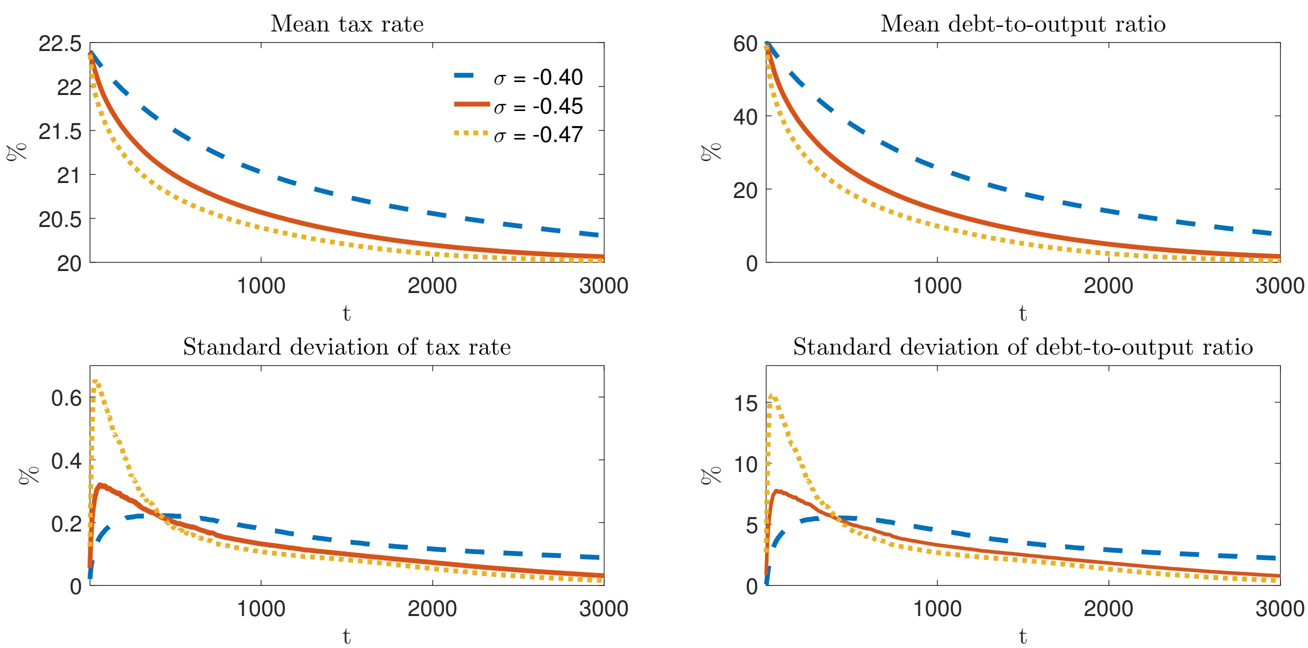

Figure 5: Moments of the tax rate and the debt-to-output ratio over time for $\sigma \in \{-0.47, -0.45, -0.40\}$. These values of σ imply a p of 29.24%, 30.21% and 32.23% respectively for $T = 50$. $10,000$ sample paths were used for each σ.

6.3 Long-run dynamics

When initial debt is positive, the Ramsey plan prescribes a deficit in the initial period. In the subsequent periods, the planner is running either a decreasing or an increasing surplus-to-output ratio, depending on the IES. In the first case, the planner repays the entire stock of debt. In the second case, the planner is postponing distortions to the future and increases public debt over time. Recall that in the full confidence case there are no drifts.

6.3.1 $IES < 1$: fiscal consolidations and long-run balanced budgets

For the baseline calibration of $\rho = 2$, we find that there is a *negative* drift with respect to π, as we expect from proposition 8. Figure 4 displays a typical sample path that captures the front-loading of distortions. Public debt converges to zero and the tax rate converges to its balanced-budget value. The intuition of this result is as follows: good times (high technology shocks) are associated with smaller taxes than bad times. Since the doubts of the household are unfounded, good times, which bear low taxes, happen more often according to π –the data-generating process– than what the pessimistic household thinks. Good times actually happen so often, so that the tax rate and public debt fall on average over time.

Doubts about the model and speed of convergence. The fiscal adjustment is initially steep and becomes flatter close to the balanced budget. Figure 5 plots the mean and standard deviation of the tax rate and the debt-to-output ratio over time for different values of σ, which imply different detection error probabilities p. The larger the doubts about the model, i.e. the lower p, the lower the mean tax rate and debt-to-output ratio and the quicker the convergence to a balanced budget. The standard deviation of the tax rate and the debt-to-output ratio behave in a non-monotonic way over time, featuring a hump-shaped pattern. The maximum standard deviation is larger for high doubts about the model. This is because the larger the doubts, the more the planner manipulates the pessimistic expectations of the agents in order to make debt cheaper and therefore the larger the changes in the tax rate and in debt, leading initially to large volatility. Then, the standard deviation of the tax rate and the debt-to-output ratio eventually decreases, till it reaches zero at the balanced budget.[24]

6.3.2 $IES > 1$: back-loading of distortions

Figure 6 displays the mean and the standard deviation of the tax rate and the debt-to-output ratio when the IES is larger than unity. As expected, there is a positive drift in the tax rate, which is reflected in the debt-to-output ratio. The intuition is similar as before: since good times happen more often according to the data-generating process than what the pessimistic household thinks,

[24] To understand the initial increase of volatility, assume for instance that the tax rate was a stationary $AR(1)$ with autocorrelation ϕ and conditional standard deviation σ_ϵ. Then, the standard deviation would increase over time till it reached its stationary counterpart, $\sigma_\epsilon/\sqrt{1-\phi^2}$. In our case of a non-stationary process that becomes eventually deterministic, after the initial increase, the standard deviation starts decreasing till it reaches zero at the balanced budget. The higher the doubts about the model, the quicker the standard deviation reaches its peak, and the quicker it approaches zero.

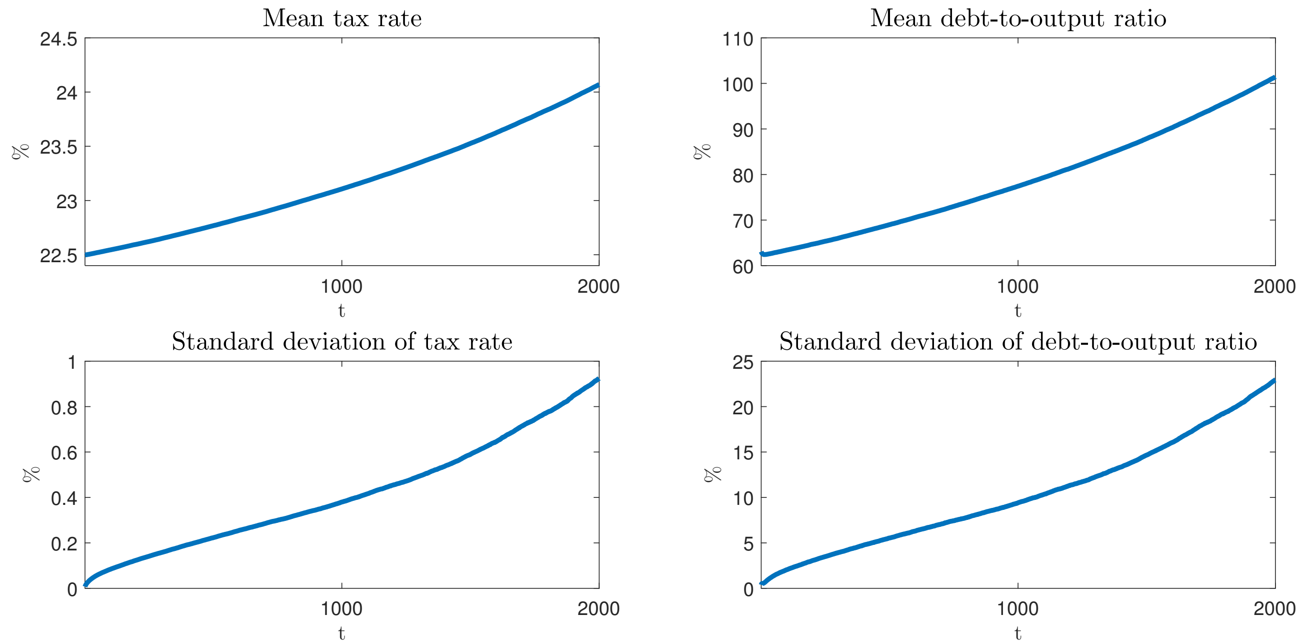

Figure 6: Mean and standard deviation of the tax rate and the debt-to-output ratio when $(\rho, \psi) = (0.5, 1)$. Doubts about the model are set to $\sigma = -1.7$, that corresponds to a detection error probability of 40.94% for $T = 50$, and a timing premium of 0.17%.

and since now good times bear higher taxes than bad times, average tax rates (and debt) increase over time.

Varying share Λ. We focused here on a constant Λ by setting $\psi = 1$. In the Online Appendix we provide the analysis of the case of substitutes ($\psi < 1$) and complements ($\psi > 1$). While the government share now varies, changes in this share are quantitatively small.[25] In addition, none of our results regarding short- and long-run properties of optimal distortions change. Thus, the exact nature of government consumption, substitute or complement to private consumption, is not key to our findings. In contrast, the fact that spending is endogenous, is. We discuss this in the next section.

7 Exogenous government spending

We have been interested in analyzing all relevant types of fiscal adjustment in an environment of ambiguity about the business cycle: taxes, government consumption and debt issuance. For that reason, we endogenized government spending by allowing it provide utility to the representative household. Typical optimal taxation studies treat government expenditures as *exogenous* and

[25] Additional shocks to the utility of government consumption, as in the work of Bachmann and Bai (2013), could potentially generate more variation in the government share.

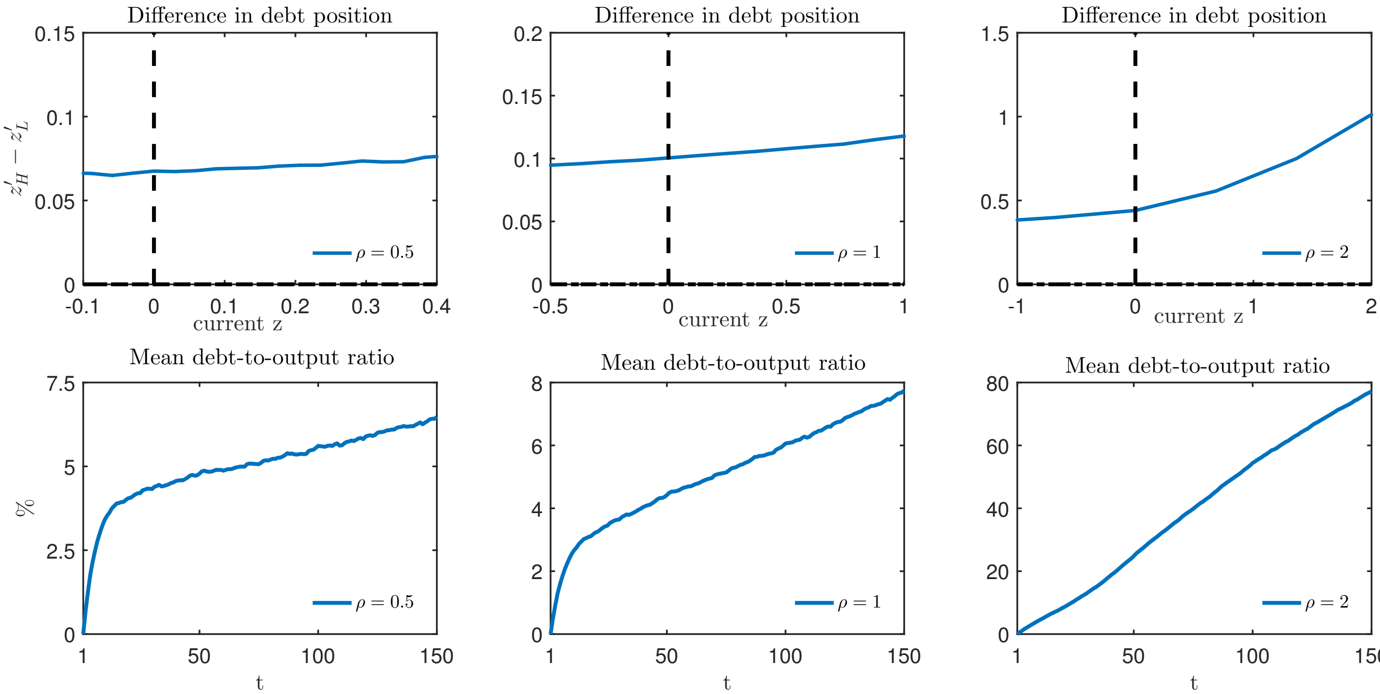

Figure 7: The top three panels from left to right depict the difference in the policy functions for next period's state-contingent debt in marginal utility units, $z'_H - z'_L$, for $\rho \in \{0.5, 1, 2\}$ respectively. The current shock is high, $s = s_H$. The bottom three panels depict from left to right the respective mean debt-to-output ratio over time for $\rho \in \{0.5, 1, 2\}$; 10,000 sample paths were used for each ρ. Doubts about the model are kept constant at $\sigma = -0.45$, and initial debt is set to $b_0 = 0$. In all cases, $z'_H > z'_L$ and debt-to-output ratios increase on average over time.

wasteful though.[26] In what follows, we show that the government consumption margin is crucial for our fiscal austerity results.

Recall that by using utility-providing government consumption and by adopting a homothetic specification we obtained a sharp characterization of the cyclicality and the dynamics of the optimal plan in terms of the IES *only*, as we showed in propositions 6 and 8 analytically and in the previous section numerically. When government expenditures are *exogenous*, the Ramsey problem is the same as in section 4.3, with the exception that g is not a control variable anymore. As a result, we get the same law of motion (25) for the excess burden of taxation and therefore, the *same* incentives for price manipulation through the worst-case beliefs of the household.

However, as we emphasized earlier, the ultimate object of interest is the response of the "value" of surpluses and debt to shocks (in a two-period and infinite horizon economy, respectively). It is natural to conjecture that, when spending is exogenous, a positive technology shock, which increases surpluses, will translate to a reduction in marginal utility and therefore, to a reduction in marginal-utility-adjusted surpluses Ω, depending on the elasticity of marginal utility, thus the IES. Nevertheless, proposition 9 below shows that the cyclicality of Ω, and therefore, the cyclicality of the excess burden, does *not* depend solely on the IES and its size relative to unity. Instead, it

[26] See Teles (2011) for a criticism of this approach.

depends also on the *size* of the surplus through the level of taxes and the government share. These features of the economy were not relevant to our homothetic endogenous spending specification.

Proposition 9. *Assume that government consumption does not provide utility and consider a period utility function with* $U_{cl} = 0$. *Then,* $\partial\Omega/\partial s > (<) 0 \iff \tau(1-\Lambda) > (<) \epsilon_{cc}(\tau - \Lambda)$.

Proof. See the Appendix. □

To get an idea about magnitudes, assume for example that $\Lambda = 20$ percent and that $\tau = 22.42$ percent, as in our benchmark model in the previous section without robustness. Then, Ω remains procyclical as long as $\epsilon_{cc} < 7.42$ and becomes countercyclical otherwise. Thus, recalling proposition 6, the excess burden of taxation would be procyclical in a two-period economy and the austerity case *irrelevant*, unless we assumed implausibly high values of ϵ_{cc}. From a different angle, assume that $\epsilon_{cc} > 1 - \Lambda$. Then Ω is procyclical if the tax rate is sufficiently small, $\tau < \epsilon_{cc}\Lambda/(\epsilon_{cc} - 1 + \Lambda)$ and countercyclical otherwise. For $(\epsilon_{cc}, \Lambda) = (2, 0.2)$ the proposition implies that if the tax rate is below 33.3 percent (so if the surplus-to-output ratio is smaller than 13.3 percent), Ω is procyclical. For larger tax rates, corresponding to very large surpluses (or debt in infinite horizon), the opposite would hold.

We dig deeper into these findings by computing global solutions to the Ramsey problem with exogenous government spending in infinite horizon. To stay close to our endogenous spending case, we use the following utility function:

$$U = \frac{c^{1-\rho} - 1}{1-\rho} - a_h \frac{h^{1+\phi_h}}{1+\phi_h}.$$

The level of g is constant and calibrated to be equal to 20 percent of output on average at the first-best. The technology shocks and the rest of the preference parameters are calibrated as in the numerical exercises section. The top panels of figure 7 display the difference in policy functions for state-contingent debt in marginal utility units for $\rho \in \{0.5, 1, 2\}$. For all parameterizations, debt in marginal utility units is procyclical, $z'_H > z'_L$, leading to a procyclical excess burden of taxation, $\Phi'_H > \Phi'_L$. Therefore, for "reasonable" calibrations, whether or not the IES is smaller, equal to, or larger than unity, taxes are procyclical, and the optimal fiscal plan exhibits a positive drift in public debt, as shown in the bottom panels of figure 7.

To conclude, utility-providing government expenditures are crucial for both the sharp dependence on the IES and for the quantitative relevance of the novel austerity result. On a more general level, these results show that the question of optimal financing of government expenditures cannot be independent of what these expenditures do in the economy and how they affect interest rates. If we think of government expenditures as war and peace shocks, as in the typical optimal taxation problem, then procyclical taxes and a positive drift in debt are optimal in environments with ambiguity about the cycle, altering non-trivially the acyclicality of taxes and lack of drifts of expected utility setups. In contrast, if we think of government expenditures as providing utility, as in the work of Bachmann and Bai (2013), then fiscal consolidations and balanced budgets can be optimal in the long-run if interest rates are sufficiently responsive, i.e. if the IES is smaller than unity.

8 Concluding remarks

In this paper we studied the optimal design of taxes, spending and pubic debt, when there is ambiguity about the cycle. We found that two, diametrically opposite, policies can be optimal: "austerity" policies, i.e. cycle-amplifying taxes and front-loaded fiscal consolidations when the IES is below unity, and "anti-austerity" policies, i.e. cycle-mitigating taxes and an increasing public debt over time when the IES is larger than unity. Typical calibrations of the IES feature values below unity, which, given this study, make the austerity case difficult to dismiss.[27]

In our study we abstracted from many other features that are in principle relevant for fiscal policy. For instance, we did not consider under-utilization of resources or any kind of default risk, that may annul or favor fiscal consolidation arguments.[28] Incorporating market incompleteness as in Bhandari et al. (2016) would be another interesting extension. Despite these limitations in scope, we find it interesting and somewhat unexpected the fact that when there are pessimistic scenarios about the economic cycle, it may actually be optimal to promote austerity measures and amplify the endogenous pessimism of the households, in order to reduce interest rates in recessions.

[27] See Guvenen (2006) and references therein for the debate on the size of the IES.

[28] Interesting work incorporating default risk and fiscal policy considerations is done by Cuadra et al. (2010), Bi (2012) and Arellano and Bai (2016).

A Full Confidence in the model

A.1 First-order conditions of second-best problem

Let Φ denote the multiplier on the unique implementability constraint and let $\beta^t \pi_t(s^t)\lambda_t(s^t)$ denote the multipliers on the resource constraint at each t, s^t. First-order necessary conditions for $t \geq 1$ are

$$c_t(s^t): \quad U_c(s^t) + \Phi\Omega_c(s^t) = \lambda_t(s^t) \tag{A.1}$$

$$h_t(s^t): \quad -U_l(s^t) + \Phi\Omega_h(s^t) = -\lambda_t(s^t)s_t \tag{A.2}$$

$$g_t(s^t): \quad U_g(s^t) + \Phi\Omega_g(s^t) = \lambda_t(s^t) \tag{A.3}$$

where $\Omega_i, i = c, h, g$ stands for the respective partial derivative of the surplus in marginal utility units Ω.

The presence of initial debt modifies the first-order conditions for $t = 0$. In particular, we have

$$c_0: \quad U_{c0} + \Phi(\Omega_{c0} - U_{cc0}b_0) = \lambda_0 \tag{A.4}$$

$$h_0: \quad -U_{l0} + \Phi(\Omega_{h0} + U_{cl0}b_0) = -\lambda_0 s_0 \tag{A.5}$$

$$g_0: \quad U_{g0} + \Phi(\Omega_{g0} - U_{cg0}b_0) = \lambda_0 \tag{A.6}$$

A.2 Proof of Proposition 1

Eliminate λ_t from (A.1), (A.2) and (A.3) to get

$$\frac{U_g + \Phi\Omega_g}{U_c + \Phi\Omega_c} = 1 \tag{A.7}$$

$$\frac{U_l - \Phi\Omega_h}{U_c + \Phi\Omega_c} = s. \tag{A.8}$$

Expressions (A.7) and (A.8) capture the optimal *wedges* at the two margins and contrast to (8) and (9) of the first-best allocation (which correspond to the case of $\Phi = 0$). Using (A.7) and (A.8) together with the resource constraint (1) allows us to solve for the optimal second-best allocation (c, h, g) in terms of the *current* technology shock s_t and the multiplier Φ, $c_t = c(s_t, \Phi), h_t = h(s_t, \Phi), g_t = g(s_t, \Phi), t \geq 1$, which proves the history-independence property. Furthermore, since the public wedge and the labor tax $\tau = 1 - U_l/(U_c s)$ are functions of the optimal allocation, they also inherit the history-independence property, $\chi_t = \chi(s_t, \Phi), \tau_t = \tau(s_t, \Phi)$.

Performing the same exercise at $t = 0$ we get

$$\frac{U_{g0}+\Phi(\Omega_{g0}-U_{cg0}b_0)}{U_{c0}+\Phi(\Omega_{c0}-U_{cc0}b_0)} = 1 \tag{A.9}$$
$$\frac{U_{l0}-\Phi(\Omega_{h0}+U_{cl0}b_0)}{U_{c0}+\Phi(\Omega_{c0}-U_{cc0}b_0)} = s_0 \tag{A.10}$$

which lead to an initial allocation that depends on (s_0, b_0, Φ). The value of the multiplier Φ is such that the implementability constraint holds, i.e. the present value of government surpluses is equal to initial debt.

A.3 Proof of Proposition 2

The public wedge and labor tax at $t = 0$ are

$$\chi_0 = \frac{\Phi(1-\epsilon_{cc}-\epsilon_{ch}-\epsilon_{gc}-\epsilon_{gh}+(\epsilon_{cc}+\epsilon_{gc})c_0^{-1}b_0)}{1+\Phi(\epsilon_{gc}+\epsilon_{gh}-\epsilon_{gc}c_0^{-1}b_0)}$$
$$\tau_0 = \frac{\Phi(\epsilon_{cc}+\epsilon_{ch}+\epsilon_{hh}+\epsilon_{hc}-(\epsilon_{cc}+\epsilon_{hc})c_0^{-1}b_0)}{1+\Phi(1+\epsilon_{hh}+\epsilon_{hc}-\epsilon_{hc}c_0^{-1}b_0)}$$

and the same comment about the positivity of the denominators applies.

To prove proposition 2 express (A.7) as $\frac{U_g}{U_c}\cdot\frac{1+\Phi\Omega_g/U_g}{1+\Phi\Omega_c/U_c} = 1$ and use the definition of the public wedge (11) to get $\chi = \frac{\Phi(\Omega_c/U_c-\Omega_g/U_g)}{1+\Phi\Omega_g/U_g}$. Similarly, express the optimal wedge in labor supply (A.8) as $\frac{U_l}{U_c}\cdot\frac{1-\Phi\Omega_g/U_l}{1+\Phi\Omega_c/U_c} = s$, which can be written in terms of the labor tax as $\tau = \frac{-\Phi(\Omega_c/U_c+\Omega_h/U_l)}{1-\Phi\Omega_h/U_l}$, since $\tau = 1-U_l/(U_c s)$. The partial derivatives of Ω scaled by the respective marginal utilities take the form $\Omega_c/U_c = 1-\epsilon_{cc}-\epsilon_{ch}$, $\Omega_h/U_l = -1-\epsilon_{hh}-\epsilon_{hc}$ and $\Omega_g/U_g = \epsilon_{gc}+\epsilon_{gh}$. Use these expressions to finally get the optimal public wedge and labor tax stated in the proposition. Use (A.9) and (A.10) and follow the same steps for $t = 0$. For the signs of the denominators, use (A.2) and (A.3) to get $1+\Phi(1+\epsilon_{hh}+\epsilon_{hc}) = \lambda s/U_l > 0$ and $1+\Phi(\epsilon_{gc}+\epsilon_{gh}) = \lambda/U_g > 0$ since $\lambda > 0$. Similarly, use (A.5) and (A.6) for $t = 0$.

A.4 Proof of Proposition 3

The marginal rate of substitution of government for private consumption is $U_g/U_c = A\kappa^{-\psi}$, where $A \equiv \alpha/(1-\alpha)$ and $\kappa \equiv g/c$, the ratio of government to private consumption. At the first-best we have $\kappa = \kappa^{FB} \equiv A^{1/\psi}$.

1a) In order to determine the optimal value of χ we need to solve the equation $U_g/U_c = 1+\chi$, which can be expressed in terms of κ as

$$A\kappa^{-\psi} = 1 + \frac{\Phi(1-\psi)}{1+\Phi(\psi-\rho)[1+A\kappa^{1-\psi}]^{-1}}. \tag{A.11}$$

This equation is derived by expressing the weight λ_c as a function of κ, $\lambda_c(\kappa) = [1 + A\kappa^{1-\psi}]^{-1}$. Equation (A.11) does not depend on the shocks s and defines implicitly κ as a function of the excess burden of taxation Φ, $\kappa(\Phi)$, with $\kappa(0)$ denoting the first-best solution. Since Φ is constant, κ and the public wedge χ become *constant* at the second-best and do not vary across states and dates. Thus, the share of government consumption in output $\Lambda \equiv \kappa/(1+\kappa)$ becomes a function of Φ, $\Lambda = \Lambda(\Phi)$, and does not vary across states and dates either.

1b) Aside from the first-best, there is no closed-form solution of (A.11) unless specific assumptions are made. For example, for $\psi = 1$ we have $\kappa(\Phi) = \kappa(0) = A$. Furthermore, if we don't differentiate between intratemporal and intertemporal substitution and set $\psi = \rho$, we get $\chi = \Phi(1-\psi)$ and $\kappa = (A/(1+\Phi(1-\psi))^{1/\psi}$. More generally, we can use the implicit function theorem to show the existence of κ and its sensitivity with respect to the excess burden of taxation. Note at first that since $\Lambda'(\Phi) = \kappa'(\Phi)/(1+\kappa)^2$, we have $\operatorname{sign}\Lambda'(\Phi) = \operatorname{sign}\kappa'(\Phi)$. Define $\mathcal{H}(\kappa,\Phi) \equiv A\kappa^{-\psi} - 1 - \Phi(1-\psi)[1+\Phi(\psi-\rho)(1+A\kappa^{1-\psi})^{-1}]^{-1}$ and write (A.11) as $\mathcal{H}(\kappa,\Phi) = 0$. By the implicit function theorem, there exists a function $\kappa(\Phi)$ in a neighborhood of a solution of the equation with derivative $\kappa'(\Phi) = -\mathcal{H}_\Phi/\mathcal{H}_\kappa$ as long as $\mathcal{H}_\kappa \neq 0$ at the solution. We have $\mathcal{H}_\Phi = (\psi-1)[1+\Phi(\psi-\rho)\lambda_c]^{-2}$ and $\mathcal{H}_\kappa = -A\kappa^{-\psi}\left[\psi\kappa^{-1} + (\psi-\rho)\Phi^2(1-\psi)^2[\lambda_c^{-1} + \Phi(\psi-\rho)]^{-2}\right]$. The sign of $\mathcal{H}_\Phi$ depends only on ψ being larger or smaller than unity, $\operatorname{sign}\mathcal{H}_\Phi = \operatorname{sign}(\psi-1)$. The partial $\mathcal{H}_\kappa$ is always negative for $\psi \geq \rho$. So for $\psi \geq \rho$ we have $\operatorname{sign}(\kappa'(\Phi)) = \operatorname{sign}(\psi-1)$. For $\psi < \rho$ the sign of $\mathcal{H}_\kappa$ is ambiguous. But it is easy to see that around the first-best solution, we have $\mathcal{H}_\kappa(\kappa^{FB},0) = -\psi/\kappa^{FB}$ and $\kappa'(0) = (\psi-1)\kappa^{FB}/\psi$. Thus, $\operatorname{sign}\Lambda'(0) = \operatorname{sign}\kappa'(0) = \operatorname{sign}(\psi-1)$.

2a) The optimal tax rate in proposition 2 becomes $\tau = \frac{\Phi(\epsilon_{cc}(\kappa)+\phi_h)}{1+\Phi(1+\phi_h)}$. The elasticity ϵ_{cc} depends on the ratio κ through the weight $\lambda_c(\kappa)$. A constant excess burden of taxation Φ leads to a constant κ and therefore ϵ_{cc} does *not* vary across shocks. Therefore, the labor tax becomes *constant* across states and dates, $\tau_t = \tau(\Phi), t \geq 1$.

2b) Differentiating the tax rate with respect to Φ delivers

$$\tau'(\Phi) = \frac{\epsilon_{cc} + \phi_h + \Phi\epsilon'_{cc}(\kappa)\kappa'(\Phi)(1+\Phi(1+\phi_h))}{(1+\Phi(1+\phi_h))^2}.$$

with $\epsilon'_{cc}(\kappa) = (\rho-\psi)(\psi-1)A\kappa^{-\psi}\lambda_c^2$. For the case of $\psi = 1$ or the $\psi = \rho$, where we have $\epsilon_{cc} = \alpha + (1-\alpha)\rho$ and $\epsilon_{cc} = \rho = \psi$ respectively, the tax rate becomes an *increasing* function of Φ.

More generally, for a small deviation from the first-best we have $\tau'(0) = \epsilon_{cc}(\kappa^{FB}) + \phi_h > 0$.

3) Assume that $b_0 = 0$. Then the initial tax rate and government share are the same as in the subsequent periods, so $\tau_t = \tau(\Phi), \Lambda_t = \Lambda(\Phi) \forall t \geq 0$. The intertemporal budget constraint reads $0 = (\tau(\Phi) - \Lambda(\Phi)) \sum_{t=0}^{\infty} \sum_{s^t} q_t(s^t) y_t(s_t)$ and therefore $\tau(\Phi) = \Lambda(\Phi)$. This equation, which is to be solved for Φ, does not depend on the shocks but only on the preference parameters $(\alpha, \rho, \psi, \phi_h)$. Thus, Φ and therefore the optimal tax rate and share Λ will not depend on stochastic properties of the shocks. When $b_0 > 0$, the intertemporal budget constraint can be rearranged to get $\tau(\Phi) - \Lambda(\Phi) = (b_0 - (\tau_0 - \Lambda_0) y_0) / \sum_{t=1}^{\infty} \sum_{s^t} q_t(s^t) y_t(s_t)$. If $b_0 > (\tau_0 - \Lambda_0) y_0$, then the government always runs surpluses $\tau(\Phi) > \Lambda(\Phi)$ for each $t \geq 1$. The value of the excess burden of taxation Φ that satisfies the budget constraint will depend on the properties of the shocks.

B Doubts about the model

B.1 Initial period problem

The recursive problem from period one onward uses as an input the value of the state variable at $t = 1$, when the shock takes value is s, $z_{1,s}$. This value is chosen *optimally* at $t = 0$, together with the initial allocation (c_0, h_0, g_0) to solve the problem

$$\max_{c_0, g_0 \geq 0, h_0 \in [0,1], z_{1,s} \in Z(s)} U(c_0, 1 - h_0, g_0) + \frac{\beta}{\sigma} \ln \sum_s \pi(s|s_0) \exp(\sigma V(z_{1,s}, s))$$

subject to

$$\begin{aligned} U_c(c_0, 1 - h_0, g_0) b_0 &= \Omega(c_0, h_0, g_0) + \beta \sum_s \pi(s|s_0) \frac{\exp(\sigma V(z_{1,s}, s))}{\sum_s \pi(s|s_0) \exp(\sigma V(z_{1,s}, s))} z_{1,s} \\ c_0 + g_0 &= s_0 h_0 \end{aligned}$$

The optimality conditions with respect to (c_0, h_0, g_0) are the same as in the problem without doubts (A.4)-(A.6), with the qualification that the multiplier on the initial implementability constraint is indexed by $t = 0$, Φ_0. Similarly, the optimality condition with respect to $z_{1,s}$ is given by (24) with the same qualification.

B.2 Proof of Proposition 5

1) We will show that, given the assumption, a constant Φ satisfies the optimality conditions of the Ramsey problem with doubts about the probability model (assuming implicitly that they are sufficient for the characterization of the solution). Debt in marginal utility units is $z_t = E_t \sum_{i=0}^{\infty} \beta^i \frac{M_{t+i}}{M_t} \Omega^\star(s_{t+i}, \Phi_{t+i})$. For any constant Φ we get $z_t = z = \Omega^\star/(1-\beta), t \geq 1$, since $\Omega^\star$ does not vary across shocks and $E_t M_{t+i} = M_t, i \geq 0$. Thus, η_{t+1} is identically zero $\forall t \geq 0$ and the law of motion for Φ_t (25) delivers $\Phi_t = \Phi, t \geq 0$, confirming that a constant Φ satisfies the optimality conditions. The constant Φ has to satisfy the implementability constraint at $t = 0$, which reduces to $U_{c0} b_0 = \Omega_0 + \beta \Omega^\star/(1-\beta)$. This is the same equation that Φ has to satisfy at the second-best with full confidence in the model. Let the solution to it be $\bar{\Phi}$ and the result follows.

2) Given the assumption, there is a $\bar{\Phi}$ for which the government runs a balanced budget for every realization of the shock (if there is more than one, we always pick the smallest one). For the given $\bar{\Phi}$ we have $z_t = 0 \forall t \geq 1$ and therefore $\eta_{t+1} \equiv 0, t \geq 0$. Thus, we have $\Phi_t = \bar{\Phi}, t \geq 0$ by (25). This $\bar{\Phi}$ satisfies the implementability constraint at $t = 0$ since initial debt is zero. This is the same condition as with full confidence in the model and the result follows. Note that it is important to have zero initial debt. If $b_0 \neq 0$, the implementability constraint would become $U_{c0} b0 = \Omega_0$. However, Ω_0 depends on $(s_0, b_0, \bar{\Phi})$ through the initial allocation (c_0, h_0, g_0) and there is no guarantee that the constraint holds for the given $\bar{\Phi}$ that furnishes a balanced budget. Other values of a constant Φ could lead to non-zero positions z_{t+1} that vary across shocks, leading to a time-varying excess burden by the law of motion (25) and annulling the conclusion.

B.3 Proof of corollary to Proposition 5

1) We will show that $\Omega^\star$ doesn't vary across shocks for any subutility of leisure $v(l)$ if $\rho = 1$. For a generic $v(l)$ the elasticity of marginal disutility of leisure (which is the inverse of the Frisch elasticity) depends on h, $\epsilon_{hh}(h) = -v''(1-h)h/v'(1-h)$, which could in principle lead to a varying tax rate across shocks for a given Φ, since $\tau = \frac{\Phi(\epsilon_{cc}(\kappa) + \epsilon_{hh}(h))}{1+\Phi(1+\epsilon_{hh}(h))}$ according to the formula in proposition 2. We will show that for $\rho = 1$, optimal labor is only a function of Φ, a fact that ultimately delivers the result. For $\rho = 1$, $U_c = \lambda_c(\kappa) c^{-1}$ and $\epsilon_{cc}(\kappa) = \psi + (1-\psi)\lambda_c(\kappa)$. Thus, the optimal wedge (A.8) becomes $\frac{v'(1-h)}{\lambda_c(\kappa)} \cdot \frac{1+\Phi(1+\epsilon_{hh}(h))}{1+\Phi(1-\epsilon_{cc}(\kappa))} c = s$. Setting $c = (1-\Lambda) sh$, leads to the elimination of the shocks s from the optimal wedge equation, furnishing a labor that is only a function of Φ. As a result, the tax rate becomes a function of only Φ (albeit a different function than in the constant Frisch case). The optimal surplus is marginal utility units becomes $\Omega^* = \lambda_c(\kappa)(\tau - \Lambda) c^{-1} y = \lambda_c(\kappa)(\tau - \Lambda)/(1-\Lambda)$, which depends only on Φ.

2) In that case, balanced budgets are optimal according to proposition 3. Therefore, $\Omega^\star(s, \bar{\Phi}) = \Omega^\star(s', \bar{\Phi}) = 0, \forall s \neq s'$, for the $\bar{\Phi}$ that satisfies $\tau(\bar{\Phi}) = \Lambda(\bar{\Phi})$.

B.4 Proof of Proposition 6

1) Express all variables in the law of motion of Φ as functions of σ to get $\Phi_i(\sigma)(1+\sigma\eta_i(\sigma)\Phi_0(\sigma)) = \Phi_0(\sigma), i = L, H$. For $\sigma = 0$ the excess burden is $\Phi(0)$ and we have $\Phi_i(0) = \Phi_0(0) = \Phi(0), i = L, H$. Let $\eta_i(0) = \Omega_i^{\sigma=0} - \sum_i \pi_i \Omega_i^{\sigma=0}, i = L, H$ denote the relative debt position for $\sigma = 0$. Differentiate with respect to σ and set $\sigma = 0$ to get $\Phi_i'(0) = \Phi_0'(0) - \Phi(0)^2\eta_i(0)$.[29] To first-order we have $\Phi_i(\sigma) \simeq \Phi(0)+\sigma\Phi_i'(0)$ and $\Phi_0(\sigma) \simeq \Phi(0)+\sigma\Phi_0'(0)$. Therefore, $\Phi_i(\sigma)-\Phi_0(\sigma) = \sigma(\Phi_i'(0)-\Phi_0'(0)) = -\sigma\Phi(0)^2\eta_i(0)$. Since $\sigma < 0$, $\Phi_H(\sigma) > \Phi_0(\sigma) > \Phi_L(\sigma)$, when $\Omega_H^{\sigma=0} > \Omega_L^{\sigma=0}$. The opposite holds when $\Omega_H^{\sigma=0} < \Omega_L^{\sigma=0}$.

2) Consider first equilibrium labor and output. Use the labor supply condition (6) and express the marginal utility of consumption as $U_c = (1-\alpha)\left(\frac{c}{u}\right)^{\rho-\psi} c^{-\rho}$ to solve for labor h and then for output, $y = sh$. We have

$$\begin{aligned} h(s,\Phi) &= H(\Phi)\cdot s^{\frac{1-\rho}{\rho+\phi_h}}, \quad \text{and} \quad y(s,\Phi) = H(\Phi)\cdot s^{\frac{1+\phi_h}{\rho+\phi_h}}, \text{ where} \\ H(\Phi) &\equiv \left[\frac{1-\tau}{a_h}\frac{(1-\alpha)\left(\frac{c}{u}\right)^{\rho-\psi}}{(1-\Lambda)^\rho}\right]^{\frac{1}{\rho+\phi_h}} \end{aligned}$$

Note that c/u is a function of κ, $c/u = \left[1-\alpha+\alpha\kappa^{1-\psi}\right]^{\frac{1}{\psi-1}}$. Therefore, H is function only of Φ, through $\tau(\Phi), \Lambda(\Phi)$ and $\kappa(\Phi)$. The income and substitution effects in labor supply are controlled only by ρ for this utility function (and not by ρ *and* ψ). The surplus is $S(s,\Phi) = (\tau(\Phi)-\Lambda(\Phi))y(s,\Phi)$. Since $\partial y/\partial s > 0$, the surplus is increasing in s for $\tau > \Lambda$. To get $\Omega^\star$, multiply S with U_c (expressed again as previously) and use $c = (1-\Lambda)y$. The expression for J is $J(\Phi) \equiv \frac{(1-\alpha)\left(\frac{c}{u}\right)^{\rho-\psi}}{(1-\Lambda)^\rho} > 0$, and is a function only of Φ (and not s), since the ratio c/u depends only on Φ. With full confidence in the model Φ is constant, and therefore $\partial\Omega^\star/\partial s$ determines the size of the surplus in marginal utility units across shocks. We have obviously sign $\partial\Omega^\star/\partial s = \text{sign}(1-\rho)$ when $\tau > \Lambda$. The result follows.

C Numerical solution method

The code which computes global solutions is divided into three parts. First, we solve the *static* problem of finding the optimal allocation (c, g, h) for a given level of surplus in marginal utility

[29] For simplicity, we use the same notation as in some parts of proposition 3, where we wanted to express small deviations from the first-best, $\Phi = 0$.

units and a given technology shock, (Ω, s). We compute the function $\mathcal{U}(\Omega, s)$, defined as:

$$\begin{aligned}\mathcal{U}(\Omega, s) = \max_{c,g,h} U(c, g, h) \text{ s.t.}\\ sh = c + g\\ \Omega = U_c c + U_h h\end{aligned}$$

We approximate this function with cubic splines for each realization of the shock s. We obtain policy functions $\mathcal{C}(\Omega, s)$, $\mathcal{G}(\Omega, s)$, $\mathcal{H}(\Omega, s)$ as the argmax of the problem stated above.

In the second step of the algorithm, we perform *value function iteration*. We solve the following problem:

$$V(z, s) = \max_{\{\Omega, z_{s'}\}} \mathcal{U}(\Omega, s) + \frac{\beta}{\sigma} \ln \sum_{s'|s} \pi(s'|s) \exp(\sigma V(s', z'_{s'})) \quad \text{s.t.} \tag{C.1}$$

$$\Omega = z - \beta \sum_{s'|s} \pi(s'|s) \frac{\exp(\sigma V(s', z'_{s'})) z'_{s'}}{\sum_{s'|s} \pi(s'|s) \exp(\sigma V(s', z'_{s'}))} \tag{C.2}$$

As an initial guess for the iteration we compute value functions as if the planner was making Φ constant over states and dates, a policy which would be suboptimal. Then, we conduct value function iteration using a simple grid search to find the optimal portfolio choice $\{z'_{s'}\}$ for each point of the state space (z, s); the value function is updated using two loops: first an inner loop, where V is only updated in (C.1), then an outer loop, in which V is also updated in (C.2). Finally, we use the value function obtained through the grid-search optimization as a first guess for a value function iteration algorithm that uses a continuous optimization routine. We approximate the value functions with cubic splines and provide also analytical derivatives to the optimization routine. We iterate until convergence to obtain $V(z, s)$. At this stage, we have also obtained policy functions $Z(z, s; s')$ and $\hat{\Omega}(z, s)$. The implied policy functions for allocations are $C(z, s) \equiv \mathcal{C}(\hat{\Omega}(z, s), s)$, $G(z, s) \equiv \mathcal{G}(\hat{\Omega}(z, s), s)$, and $H(z, s) \equiv \mathcal{H}(\hat{\Omega}(z, s), s)$.

After solving for the value function that represents the value of the commitment problem from $t = 1$ onward, we turn to the solution of the *time-zero problem*. Given the initial conditions (s_0, b_0), we find the optimal allocation (c_0, g_0, h_0) and the optimal initial value of the pseudo-state variable, z_{1,s_1}, that maximizes the utility of the household at $t = 0$.

D Timing premium

In Epstein et al. (2014) the timing premium is defined as the fraction of the current and future consumption stream that the decision maker would be willing to give up for all risk to be resolved at $t = 1$. The decision maker faces an exogenous stochastic stream of consumption. Utility of the consumption stream when uncertainty is resolved gradually is compared to utility obtained at

$t=0$ when uncertainty is resolved at $t=1$, that is, when all future shocks, and therefore all future allocations, are known at $t=1$.

In the context of an optimal policy problem, the definition of the timing premium is more involved. In particular, we allow the planner to choose *optimally* policy for the case when uncertainty is resolved at $t=1$. Our planner faces, say, N deterministic paths at $t=1$, that are random from the perspective of $t=0$. We retain the complete market assumption by allowing the planner to issue at $t=0$ debt contingent on these N paths.

The algorithm is as follows. Let V_0 denote the utility when uncertainty is resolved gradually. We compute optimal policy when risk is resolved at $t=1$, using a fixed time horizon of $T=1,000$ years (pasting $V(z,s)$ as the continuation value at T) and $N=1,000$ simulations.[30] For each history n, we obtain an initial allocation $\{c_0^{rr}, h_0^{rr}, g_0^{rr}\}$, allocations from $t=1$ up to $t=T-1$ $\{\{c_t^{rr}(n), h_t^{rr}(n), g_t^{rr}(n)\}_{n=1}^{N}\}_{t=1}^{T-1}$ and a final debt position $\{z_T^{rr}(n)\}_{n=1}^{N}$. Finally, we compute the fraction π^{timing} that the decision-maker would be willing to give up such that:

$$W^{rr}(\pi^{\text{timing}}) = V_0$$

where $W^{rr}(\pi^{\text{timing}})$ is the utility at $t=0$ under the scenario of early resolution of risk when the decision maker gives up a fraction π^{timing} of the consumption stream. This number is computed in two steps. For each possible history n we compute

$$W_n(\pi^{\text{timing}}) \equiv \sum_{t=1}^{T-1} \beta^{t-1} U((1-\pi^{\text{timing}})c_t^{rr}(n), 1-h_t^{rr}(n), g_t^{rr}(n)) + \beta^{T-1} V(z_T^{rr}(n), s_T^{rr}(n)) \; \forall n,$$

where $s_T^{rr}(n)$ stands for the realization of the shock at period T for history n. At $t=0$, we have

$$W^{rr}(\pi^{\text{timing}}) \equiv U((1-\pi^{\text{timing}})c_0^{rr}, 1-h_0^{rr}, g_0^{rr}) + \frac{\beta}{\sigma} \ln \sum_{n=1}^{N} \frac{1}{N} \exp(\sigma W_n(\pi^{\text{timing}})).$$

Another avenue we could follow would be to treat consumption, hours worked, and government spending as *exogenous* stochastic variables, in order to be closer to the spirit of Epstein et al. (2014). Such a treatment of the allocations captures the pure desire for early resolution of *consumption* (and leisure and government consumption) uncertainty, so no part of the timing premium can be attributed to any kind of planning advantage due to the early resolution of the inherent uncertainty that drives the economy– the technology shocks in our case. In the context of the calculation above, for each history n we would use the allocation that was found to be optimal for the same history of shock realizations when uncertainty is resolved gradually – which is obviously a suboptimal allocation given the new specification of uncertainty. This way of approaching the problem would

[30] We are restricted computationally in the size of N given the *optimal* choice of N history-contingent contracts at $t=0$.

lead to a smaller π^{timing}. We conducted this exercise as well and found that numbers were virtually identical.

E Proof of Proposition 9

Assume that government expenditures are exogenous and constant. The system

$$\begin{aligned} U_l(c, 1-h) &= U_c(c, 1-h)(1-\tau)s \\ c + g &= sh = y \end{aligned}$$

determines implicitly the allocation (c, h) as function of (s, τ). Differentiate with respect to s to get the system

$$\begin{pmatrix} U_{cl} - s(1-\tau)U_{cc} & s(1-\tau)U_{cl} - U_{ll} \\ 1 & -s \end{pmatrix} \begin{pmatrix} \frac{\partial c}{\partial s} \\ \frac{\partial h}{\partial s} \end{pmatrix} = \begin{pmatrix} (1-\tau)U_c \\ h \end{pmatrix}$$

Assume $U_{cl} \geq 0$. The determinant of the system is $\Delta = s^2(1-\tau)U_{cc} - s(2-\tau)U_{cl} + U_{ll} < 0$. Then, we have

$$\begin{aligned} \frac{\partial c}{\partial s} &= \frac{\partial y}{\partial s} = \frac{-s(1-\tau)(U_c + U_{cl}h) + U_{ll}h}{\Delta} > 0 \\ \frac{\partial h}{\partial s} &= \frac{U_{cl}h - s(1-\tau)U_{cc}h - (1-\tau)U_c}{\Delta} \end{aligned}$$

The sign of $\partial h/\partial s$ is ambiguous and depends on the strength of income and substitution effects.

Consider now the surplus. We have $S(s, \tau) = \tau sh(s, \tau) - g$, with $\partial S/\partial s = \tau \partial y/\partial s > 0$. The surplus in MU units is $\Omega(s, \tau) \equiv U_c(c, 1-h)[\tau sh - g]$. Differentiating with respect to s we get:

$$\frac{\partial \Omega}{\partial s} = (U_{cc}\frac{\partial c}{\partial s} - U_{cl}\frac{\partial h}{\partial s})[\tau y - g] + U_c \tau \frac{\partial y}{\partial s}$$

The second term is always positive since it depicts the increase in surplus due to an increase in output. The first term can be negative if $S > 0$ due to decreasing marginal utility. Assume $U_{cl} = 0$ and use the fact that $\partial c/\partial s = \partial y/\partial s$ and that $y/c = 1/(1-\Lambda)$, to get:

$$\begin{aligned} \frac{\partial \Omega}{\partial s} &= \frac{\partial y}{\partial s} U_c \big[\frac{U_{cc}}{U_c}[\tau y - g] + \tau\big] \\ &= \frac{\partial y}{\partial s} U_c \big[-\epsilon_{cc}\frac{\tau - \Lambda}{1-\Lambda} + \tau\big]. \end{aligned}$$

The result follows.

References

Alesina, Alberto and Roberto Perotti. 1995. Fiscal expansions and adjustments in OECD economies. *Economic Policy* 21:207–48.

Ambler, Steve and Alain Paquet. 1996. Fiscal spending shocks, endogenous government spending, and real business cycles. *Journal of Economic Dynamics and Control* 20:237–256.

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association* 1 (1):68–123.

Arellano, Cristina and Yan Bai. 2016. Fiscal Austerity during Debt Crises. *Economic Theory* 1–17.

Atkinson, A. B. and J.E. Stiglitz. 1972. The structure of indirect taxation and economic efficiency. *Journal of Public Economics* 1:97–119.

Bachmann, Ruediger and Jinhui H. Bai. 2013. Public consumption over the business cycle. *Quantitative Economics* 4 (3):417–451.

Barillas, Francisco, Lars Peter Hansen, and Thomas J. Sargent. 2009. Doubts or Variability? *Journal of Economic Theory* 144 (6):2388–2418.

Barlevy, Gadi. 2009. Policymaking under Uncertainty: Gradualism and Robustness. *Economic Pespectives* Q II:38–55. Federal Reserve Bank of Chicago.

———. 2011. Robustness and Macroeconomic Policy. *Annual Review of Economics* 3:1–24.

Benigno, Pierpaolo and Salvatore Nisticò. 2012. International portfolio allocation under model uncertainty. *American Economic Journal: Macroeconomics* 4 (1):144–89.

Benigno, Pierpaolo and Luigi Paciello. 2014. Monetary policy, doubts and asset prices. *Journal of Monetary Economics* 64:85–98.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent. 2016. Fiscal Policy and Debt Management with Incomplete Markets. *Quarterly Journal of Economics* Forthcoming.

Bi, Huixin. 2012. Sovereign Default Risk Premia, Fiscal Limits and Fiscal Policy. *European Economic Review* 56 (3).

Bi, Huixin, Eric M. Leeper, and Campbell Leith. 2013. Uncertain Fiscal Consolidations. *Economic Journal* 123 (566).

Bianchi, Francesco, Cosmin Ilut, and Martin Schneider. 2017. Uncertainty shocks, asset supply and pricing over the business cycle. Forthcoming, Review of Economic Studies.

Bidder, Rhys M. and Matthew E. Smith. 2012. Robust animal spirits. *Journal of Monetary Economics* 59 (8):738–750.

Boyd, John H.III. 1990. Recursive utility and the Ramsey problem. *Journal of Economic Theory* 50.

Croce, Mariano M., Thien T. Nguyen, and Lukas Schmid. 2012. The market price of fiscal uncertainty. *Journal of Monetary Economics* 59 (5):401–416.

Cuadra, Gabriel, Juan M. Sanchez, and Horacio Sapriza. 2010. Fiscal policy and default risk in emerging markets. *Review of Economic Dynamics* 13:452–469.

Debortoli, Davide and Ricardo Nunes. 2013. Lack of commitment and the level of debt. *Journal of the European Economic Association* 11 (5):1053–1078.

Dovis, Alessandro, Mikhail Golosov, and Ali Shourideh. 2016. Political Economy of Sovereign Debt: A Theory of Cycles of Populism and Austerity. Mimeo, Princeton University.

Epstein, Larry G. and Martin Schneider. 2010. Ambiguity and asset markets. *Annual Reviews of Financial Economics* 2:315–334.

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.

Guvenen, Fatih. 2006. Reconciling conflicting evidence on the elasticity of intertemporal substitution: A macroeconomic perspective. *Journal of Monetary Economics* 53 (7):1451–1472.

Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.

———. 2008. *Robustness.* Princeton University Press.

Ilut, Cosmin, Rosen Valchev, and Nicolas Vincent. 2016. Paralyzed by Fear: Rid and Discrete Pricing under Demand Uncertainty. Mimeo, Duke University.

Ilut, Cosmin L. and Martin Schneider. 2014. Ambiguous business cycles. *American Economic Review* 104 (8):2368–99.

Karantounias, Anastasios G. 2013a. Managing pessimistic expectations and fiscal policy. *Theoretical Economics* 8 (1):193–231.

———. 2013b. Optimal fiscal policy with recursive preferences. Federal Reserve Bank of Atlanta WP 2013-7.

Klein, Paul, Per Krusell, and José Víctor Ríos Rull. 2008. Time-consistent public policy. *Review of Economic Studies* 75 (3):789–808.

Koopmans, Tjalling C. 1960. Stationary ordinal utility and impatience. *Econometrica* 28 (2):287–309.

Kopecky, Karen A. and Richard M.H. Suen. 2010. Finite state Markov-chain approximations to highly persistent processes. *Review of Economic Dynamics* 13 (3):701–714.

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.

Ni, Shawn. 1995. An empirical analysis on the substitutability between private consumption and government purchases. *Journal of Monetary Economics* 36 (3):593–605.

Pouzo, Demian and Ignacio Presno. 2016. Sovereign Default Risk and Uncertainty Premia. *American Economic Journal: Macroeconomics* 8 (3):230–66.

Romei, Federica. 2014. Need for the (right) speed: the timing and composition of public debt deleveraging. Mimeo, Stockholm School of Economics.

Stockman, David R. 2001. Balanced-Budget Rules: Welfare Loss and Optimal Policies. *Review of Economic Dynamics* 4 (2):438–459.

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.

Teles, Pedro. 2011. Exogenous public spending in the optimal taxation problem. Mimeo, Banco de Portugal and Universidade Catolica Portuguesa.

Fiscal austerity in ambiguous times *

Axelle Ferriere[†] Anastasios G. Karantounias[‡]

January 17, 2018

Online Appendix

(not for publication)

Keywords: Public consumption, intertemporal elasticity of substitution, balanced budget, austerity, fiscal consolidation, ambiguity aversion, multiplier preferences.

JEL classification: D80; E62; H21; H63.

*The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

[†]Department of Economics, European University Institute, Via delle Fontanelle 18, 50014, Fiesole, Italy. Email: *axelle.ferriere@eui.eu*.

[‡]Corresponding author. Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St NE, Atlanta, GA 30309, USA. Email: *anastasios.karantounias@atl.frb.org*.

A Balanced budget approximation

A.1 Preliminaries

We proceed with an approximation around the balanced budget by treating Φ as a state variable. Let Φ^* denote the value of the excess burden that leads to a balanced budget, $\tau(\Phi^*) = \Lambda(\Phi^*)$. Whenever necessary, we use the asterisk * to denote the evaluation of a function at Φ^*. Assume shocks take N values and that they are ranked as $s_1 < s_2 < ... < s_N$. To ease notation, we let $\Omega_i(\Phi)$, $z_i(\Phi)$ and $U_i(\Phi), V_i(\Phi)$ denote the level of surplus and debt (in MU units), together with the period and discounted value of utility when the excess burden of taxation is Φ and the shock is $s_t = s_i$.[1] At the balanced budget we have obviously $\Omega_i(\Phi^*) = z_i(\Phi^*) = 0, \forall i$. Since Φ^* is an absorbing state, we can also calculate $V_i(\Phi^*)$ from the recursion

$$V_i(\Phi^*) = U_i(\Phi^*) + \frac{\beta}{\sigma} \ln \sum_j \pi(j|i) \exp(\sigma V_j(\Phi^*)), \forall i, \tag{1}$$

which delivers the respective conditional distortions $m^*_{j|i}$ at Φ^*. The matrix of distortions and the distorted transition matrix are defined respectively as

$$\mathbf{M} \equiv \begin{pmatrix} m^*_{1|1} & ... & m^*_{N|1} \\ & & \\ m^*_{1|N} & ... & m^*_{N|N} \end{pmatrix}, \quad \Pi^* \equiv \Pi \circ \mathbf{M},$$

where $\circ$ denotes element-by-element multiplication. Furthermore, we collect the derivatives of the excess burden of taxation in the $N \times N$ matrix

$$\boldsymbol{\Phi} \equiv \begin{pmatrix} \Phi'_{1|1}(\Phi^*) & ... & \Phi'_{N|1}(\Phi^*) \\ \vdots & & \\ \Phi'_{1|N}(\Phi^*) & ... & \Phi'_{N|N}(\Phi^*) \end{pmatrix}$$

A.2 Approximate law of motion

Recall that the approximate law of motion of the excess burden takes the form

$$\Phi_{j|i}(\Phi) \simeq \Phi^* + \Phi'_{j|i}(\Phi^*)(\Phi - \Phi^*), \quad i, j = 1, ..., N. \tag{2}$$

To find the entries of $\boldsymbol{\Phi}$ proceed as follows. Let the current shock be i and the current excess

[1]For simplicity we do not differentiate our notation in this section and still use Ω for the indirect function of s, Φ. So Ω_i does *not* stand anymore for the derivative of Ω with respect to c, h, g.

burden of taxation Φ. Let Φ_j denote the excess burden of taxation next period at shock j. Define

$$F_{j|i}(\Phi_1, \Phi_2, ..., \Phi_N, \Phi) \equiv \Phi_j \left[1 + \sigma \eta_{j|i}(\Phi_1, \Phi_2, ..., \Phi_N)\Phi\right] - \Phi, \forall j.$$

where

$$\begin{aligned}
\eta_{j|i}(\Phi_1, \Phi_2, ..., \Phi_N) &\equiv z_j(\Phi_j) - \sum_k \pi(k|i) m_{k|i}(\Phi_1, \Phi_2, ..., \Phi_N) z_k(\Phi_k), \forall j \\
m_{j|i}(\Phi_1, \Phi_2, ..., \Phi_N) &\equiv \frac{\exp(\sigma V_j(\Phi_j))}{\sum_k \pi(k|i) \exp(\sigma V_k(\Phi_k))}, \forall j
\end{aligned}$$

Define the vector function $\mathbf{F_i} \equiv [F_{1|i}, ..., F_{N|i}]^T, \forall i$, where T denotes transpose. Given the current shock i, the law of motion for the inverse of the excess burden of taxation implies the system $\mathbf{F_i} = \mathbf{0}$, where $\mathbf{0}$ is the $N \times 1$ zero vector. Apply the implicit function theorem at $\Phi_i = \Phi = \Phi^*, \forall i$ to get the coefficients $\Phi'_{j|i}(\Phi^*)$ of the approximate law of motion (2). In particular, we have N systems

$$J_i^* \begin{pmatrix} \Phi'_{1|i}(\Phi^*) \\ \vdots \\ \Phi'_{N|i}(\Phi^*) \end{pmatrix} = -\frac{\partial \mathbf{F_i}^*}{\partial \Phi}, \quad \forall i,$$

where J_i^* the Jacobian of $\mathbf{F_i}$ evaluated at Φ^*,

$$J_i^* \equiv \begin{pmatrix} \frac{\partial F^*_{1|i}}{\partial \Phi_1} & \cdots & \frac{\partial F^*_{1|i}}{\partial \Phi_N} \\ & & \\ \frac{\partial F^*_{N|i}}{\partial \Phi_1} & \cdots & \frac{\partial F^*_{N|i}}{\partial \Phi_N} \end{pmatrix}.$$

Derivatives of the system. The derivatives of the functions $F_{j|i}$ are

$$\begin{aligned}
&\frac{\partial F_{j|i}}{\partial \Phi} = \sigma \eta_{j|i}(\Phi_1, ..., \Phi_N)\Phi_j - 1 \Rightarrow \frac{\partial F^*_{j|i}}{\partial \Phi} = -1 \\
&\frac{\partial F_{j|i}}{\partial \Phi_j} = 1 + \sigma \eta_{j|i}(\Phi_1, ..., \Phi_N)\Phi + \sigma \Phi_j \Phi \frac{\partial \eta_{j|i}}{\partial \Phi_j} \Rightarrow \frac{\partial F^*_{j|i}}{\partial \Phi_j} = 1 + \sigma (\Phi^*)^2 \frac{\partial \eta^*_{j|i}}{\partial \Phi_j} \\
&\frac{\partial F_{j|i}}{\partial \Phi_k} = \sigma \Phi_j \Phi \partial \frac{\eta_{j|i}}{\partial \Phi_k}, k \neq j \Rightarrow \frac{\partial F^*_{j|i}}{\partial \Phi_k} = \sigma (\Phi^*)^2 \frac{\partial \eta^*_{j|i}}{\partial \Phi_k}, k \neq j
\end{aligned}$$

The simplifications at Φ^* are coming from the fact that the relative debt positions are equal to zero, $\eta^*_{j|i} = 0, \forall i, j$. So we have

$$\frac{\partial \mathbf{F_i}^*}{\partial \Phi} = -\mathbf{1} \quad \text{and} \quad J^*_i = I + \sigma \cdot (\Phi^*)^2 J^*_{\eta_\mathbf{i}},$$

where $\mathbf{1}$ the $N \times 1$ unit vector, I the identity matrix and $J^*_{\eta_\mathbf{i}}$ the Jacobian of the vector of the relative debt positions $\eta_\mathbf{i} \equiv [\eta_{1|i}, ..., \eta_{N,i}]^T$, evaluated at Φ^*. Thus, the i-th system becomes

$$\left[I + \sigma \cdot (\Phi^*)^2 J^*_{\eta_\mathbf{i}}\right] \cdot \begin{pmatrix} \Phi'_{1|i}(\Phi^*) \\ \vdots \\ \Phi'_{N|i}(\Phi^*) \end{pmatrix} = \mathbf{1}, \quad \forall i. \tag{3}$$

Derivatives of the relative debt position. Consider now the matrix $J^*_{\eta_\mathbf{i}}$. The derivatives of the relative debt positions $\eta_{j|i}$ are

$$\frac{\partial \eta_{j|i}}{\partial \Phi_j} = z'_j(\Phi_j) - \big[\sum_k \pi(k|i)\frac{\partial m_{k|i}}{\partial \Phi_j} z_k(\Phi_k) + \pi(j|i) m_{j|i} z'_j(\Phi_j)\big] \quad \Rightarrow \frac{\partial \eta^*_{j|i}}{\partial \Phi_j} = (1 - \pi(j|i) m^*_{j|i}) z'_j(\Phi^*)$$

$$\frac{\partial \eta_{j|i}}{\partial \Phi_l} = -\sum_k \pi(k|i)\frac{\partial m_{k|i}}{\partial \Phi_l} z_k(\Phi_k) - \pi(l|i) m_{l|i} z'_l(\Phi_l), l \neq j \quad \Rightarrow \frac{\partial \eta^*_{j|i}}{\partial \Phi_l} = -\pi(l|i) m^*_{l|i} z'_l(\Phi^*), l \neq j$$

Thus, the Jacobian of $\eta_\mathbf{i}$ takes the form

$$\begin{aligned} J_{\eta_\mathbf{i}^*} &= \begin{pmatrix} \left[1 - \pi(1|i) m^*_{1|i}\right] z'_1(\Phi^*) & -\pi(2|i) m^*_{2|i} z'_2(\Phi^*) & \ldots & -\pi(N|i) m^*_{N|i} z'_N(\Phi^*) \\ -\pi(1|i) m^*_{1|i} z'_1(\Phi^*) & \left[1 - \pi(2|i) m^*_{2|i}\right] z'_2(\Phi^*) & \ldots & -\pi(N|i) m^*_{N|i} z'_N(\Phi^*) \\ \\ -\pi(1|i) m^*_{1|i} z'_1(\Phi^*) & -\pi(2|i) m^*_{2|i} z'_2(\Phi^*) & \ldots & \left[1 - \pi(N|i) m^*_{N|i}\right] z'_N(\Phi^*) \end{pmatrix} \\ &= \left[I - \mathbf{1} \cdot (e_i^T \Pi^*)\right] \operatorname{diag}\{\mathbf{z'}\}, \end{aligned} \tag{4}$$

where diag denotes a diagonal matrix with the vector $\mathbf{z'} \equiv [z'_1(\Phi^*), ..., z'_N(\Phi^*)]^T$ on the diagonal. Thus, in order to solve the system (3), we need the sensitivity of the debt positions with respect to the excess burden of taxation $\mathbf{z'}$.

We are going to work under the following assumption.

Assumption 1. *Doubts about the model are such so that*

$$1 + \sigma(\Phi^*)^2 \max_i z'_i(\Phi^*) > 0 \tag{5}$$

This assumption imposes bounds on the doubts about the model if $\max_i z_i'(\Phi^*) > 0$, since in that case σ has to be small enough in absolute value, $\sigma > -1/((\Phi^*)^2 \max_i z_i'(\Phi^*))$. The restriction is implicit, in the sense that (5) depends on endogenous objects, which themselves depend on σ. It was always holding for the σ that we considered numerically.

A.3 Three lemmata

Lemma 1. *The excess burden of taxation is a martingale with respect to the worst-case transition matrix* Π^* *at a first-order approximation around* Φ^**.*

Proof. We will show that

$$\sum_j \pi(j|i) m^*_{j|i} \Phi'_{j|i}(\Phi^*) = 1, \forall i \tag{6}$$

If (6) holds, then the approximate law of motion (2) implies that $\sum_j \pi(j|i) m^*_{j|i} \Phi_{j|i}(\Phi) = \Phi$ and the result follows. To show (6) remember that the relative debt positions add to zero according to the worst-case model,

$$\sum_j \pi(j|i) m_{j|i}(\Phi_{1|i}(\Phi), ..., \Phi_{N|i}(\Phi)) \eta_{j|i}(\Phi_{1|i}(\Phi), ..., \Phi_{N|i}(\Phi)) = 0, \forall i.$$

Differentiate implicitly with respect to Φ to get

$$\sum_j \pi(j|i) \big[\sum_k \frac{\partial m j|i}{\partial \Phi_k} \Phi'_{k|i}(\Phi)\big] \eta_{j|i} + \sum_j \pi(j|i) m_{j|i} \big[\sum_k \frac{\partial \eta_{j|i}}{\partial \Phi_k} \Phi'_{k|i}(\Phi)\big] = 0$$

At Φ^* this expression simplifies to

$$\sum_j \pi(j|i) m^*_{j|i} \big[\sum_k \frac{\partial \eta^*_{j|i}}{\partial \Phi_k} \Phi'_{k|i}(\Phi^*)\big] = 0, \quad \text{or} \quad e_i^T \Pi^* J^*_{\eta_i} \cdot \begin{pmatrix} \Phi'_{1|i}(\Phi^*) \\ \vdots \\ \Phi'_{N|i}(\Phi^*) \end{pmatrix} = 0, \forall i, \tag{7}$$

where e_i the vector with unity at position i and zero otherwise. Pre-multiply system (3) with $e_i^T \Pi^*$ to get

$$e_i^T \Pi^* \cdot \begin{pmatrix} \Phi'_{1|i}(\Phi^*) \\ \vdots \\ \Phi'_{N|i}(\Phi^*) \end{pmatrix} + \sigma \cdot (\Phi^*)^2 e_i^T \Pi^* J^*_{\eta_i} \cdot \begin{pmatrix} \Phi'_{1|i}(\Phi^*) \\ \vdots \\ \Phi'_{N|i}(\Phi^*) \end{pmatrix} = e_i^T \Pi^* \cdot \mathbf{1} = 1.$$

The second term at the left-hand side above is by (7) *zero*, a fact which delivers ultimately (6). □

Lemma 2. *Assume assumption 1 holds. We have*

$$\Phi'_{j|i}(\Phi^*) = \frac{1 + \sigma(\Phi^*)^2 \sum_j \pi(j|i) m^*_{j|i} \Phi'_{j|i}(\Phi^*) z'_j(\Phi^*)}{1 + \sigma(\Phi^*)^2 z'_j(\Phi^*)}, \forall i, j. \tag{8}$$

Therefore:

- *If* $\sigma = 0$, $\Phi'_{j|i}(\Phi^*) = 1, \forall i, j$.
- *More generally, we have* $\Phi'_{j|i}(\Phi^*) > 0$*, so (6) implies that* $\mathbf{A} \equiv \Pi \circ M \circ \mathbf{\Phi}$ *is a stochastic matrix.*
- *If there is* <u>*no*</u> *variation in the derivatives of debt, i.e.* $z'_j(\Phi^*) = z'_i(\Phi^*) \forall i, j$*, then* $\Phi'_{j|i}(\Phi^*) = 1, \forall i, j$*, so* $\Phi_{j|i}(\Phi) = \Phi \forall i, j$.
- *If* $z'_k(\Phi^*) > z'_l(\Phi^*)$ *then* $\Phi'_{k|i}(\Phi^*) > \Phi'_{l|i}(\Phi^*)$.
- *If* $z'_j(\Phi^*) > (<) \sum_j \pi(j|i) m^*_{j|i} \Phi'_{j|i}(\Phi^*) z'_j(\Phi^*)$ *then* $\Phi'_{j|i}(\Phi^*) > (<) 1$.

Proof. Use the expression for $J^*_{\eta_i}$ in system (3) to get

$$\Big[I + \sigma \cdot (\Phi^*)^2 \big[I - \mathbf{1} \cdot (e_i^T \Pi^*) \big] \operatorname{diag} \{\mathbf{z}'\} \Big] \cdot \mathbf{\Phi}^T e_i = \mathbf{1}$$

Rewrite the above as

$$[I + \sigma \cdot (\Phi^*)^2 \operatorname{diag} \{\mathbf{z}'\}] \mathbf{\Phi}^T e_i = \mathbf{1}(1 + \sigma \cdot (\Phi^*)^2 e_i^T \Pi^* \operatorname{diag} \{\mathbf{z}'\} \mathbf{\Phi}^T e_i),$$

The matrix that premultiplies the left-hand side is the sum of two diagonal so it is also diagonal. We can express the system above as

$$\mathbf{\Phi}^T e_i = \operatorname{diag} \{\mathbf{1} + \sigma(\Phi^*)^2 \mathbf{z}'\}^{-1} \mathbf{1}(1 + \sigma \cdot (\Phi^*)^2 e_i^T \Pi^* \operatorname{diag} \{\mathbf{z}'\} \mathbf{\Phi}^T e_i). \tag{9}$$

The inverse of the diagonal matrix is

$$\operatorname{diag}\left\{\mathbf{1}+\sigma(\Phi^*)^2\mathbf{z}'\right\}^{-1} = \begin{pmatrix} \frac{1}{1+\sigma(\Phi^*)^2 z_1'(\Phi^*)} & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & \frac{1}{1+\sigma(\Phi^*)^2 z_N'(\Phi^*)} \end{pmatrix}$$

Furthermore, $e_i^T \Pi^* \operatorname{diag}\{\mathbf{z}'\}\mathbf{\Phi}^T e_i = \sum_j \pi(j|i) m^*_{j|i} \Phi'_{j|i}(\Phi^*) z_j'(\Phi^*)$. Thus, writing explicitly system (9) delivers (8). For $\sigma = 0$ the result is obvious from (8). Furthemore, assumption 1 implies that $1+\sigma(\Phi^*)^2 z_i'(\Phi^*) > 0 \forall i$. Therefore, (8) implies that $\Phi'_{j|i}(\Phi^*) > 0$. If there is no variation in the sensitivity of the debt positions with respect to the excess burden of taxation, then formula (8), implies again $\Phi'_{j|i}(\Phi^*) = 1$, by using (6). Furthermore, the same formula implies that the monotonicity of the entries of each row i of the matrix $\mathbf{\Phi}$, and therefore the allocation of tax distortions across shocks, depends on the monotonicity of the sensitivity of the debt positions, $z_j'(\Phi^*)$. The same comment applies for the sensitivity of the debt positions and the size of the $\Phi'_{j|i}(\Phi^*)$ with respect to unity.

□

Formula (8) connects the allocation of distortions across states and states to the *sensitivity* of the debt positions in the proximity of the balanced budget, $z_j'(\Phi^*)$. The *relative* debt sensitivity, i.e. the sensitivity at j relative to the "average" sensitivity (average according to the probability measure encoded in matrix $\Pi \circ \mathbf{M} \circ \mathbf{\Phi}$), determines the increase or decrease of the excess burden over time and states. Formula (8) provides also the direct analogue to the results of proposition **5** in the text: if there is no variation of the sensitivity of debt positions across shocks, then there is no room for price manipulation through the worst-case beliefs, and therefore no reason to vary the excess burden across states and dates.

Lemma 3. *The sensitivity of debt positions depends on the sensitivity of surplus in marginal utility units through the present discounted value formula:*

$$\mathbf{z}' = \Big(I - \beta(\Pi^* \circ \mathbf{\Phi})\Big)^{-1} \mathbf{\Omega}', \tag{10}$$

where $\mathbf{\Omega}' \equiv [\Omega_1'(\Phi^*), ..., \Omega_N'(\Phi^*)]^T$*, i.e. the vector that collects the sensitivity of the surplus in marginal utility units,* $\Omega_i'(\Phi^*)$*.*

Proof. Consider the implementability constraints

$$z_i(\Phi) = \Omega_i(\Phi) + \beta \sum_j \pi(j|i) m_{j|i}(\Phi_{1|i}(\Phi), ..., \Phi_{N|i}(\Phi)) z_j(\Phi_{j|i}(\Phi)), \forall i$$

Differentiate implicitly with respect to Φ to get

$$z_i'(\Phi) = \Omega_i'(\Phi) + \beta \sum_j \pi(j|i) \big[\sum_k \frac{\partial m_{j|i}}{\partial \Phi_k} \Phi_{k|i}'(\Phi) \big] z_j(\Phi_j) + \beta \sum_j \pi(j|i) m_{j|i} z_j'(\Phi_j) \Phi_{j|i}'(\Phi)$$

which at Φ^* becomes

$$z_i'(\Phi^*) = \Omega_i'(\Phi^*) + \beta \sum_j \pi(j|i) m_{j|i}^* \Phi_{j|i}'(\Phi^*) z_j'(\Phi^*) \forall i$$

The differentiated implementability constraints can be written as a system, $\mathbf{z}' = \mathbf{\Omega}' + \beta(\Pi^* \circ \mathbf{\Phi})\mathbf{z}'$, and the result follows.

$\square$

A.4 Proof of proposition 7

Part **1** is a direct consequence of the approximate law of motion (2). Part **2** is proved in lemmata 1 and 2. To prove part **3** note that under the assumption of decreasing $m_{j|i}^*$ in j, the reference model first-order stochastically dominates the worst-case model. Then, when $\Phi_{j|i}'(\Phi^*)$ is increasing in j, i.e. if the derivatives are increasing functions of the shock, we have $\sum_j \pi(j|i)\Phi_{j|i}'(\Phi^*) > \sum_j \pi(j|i) m_{j|i}^* \Phi_{j|i}'(\Phi^*) = 1$, where the first inequality comes from the properties of first-order stochastic dominance and the second equality from lemma 1. The opposite inequality holds if $\Phi_{j|i}'(\Phi^*)$ is decreasing in j. Use the approximate law of motion (2) to get the corresponding positive and negative drifts when $\Phi > \Phi^*$.

A.5 Proof of proposition 8

Write the surplus in marginal utility units as

$$\Omega_i(\Phi) = U_c(i, \Phi)(\tau(\Phi) - \Lambda(\Phi))y(i, \Phi)$$

Differentiating with respect to Φ and evaluating at Φ^* delivers

$$\Omega_i'(\Phi^*) = \big(\tau'(\Phi^*) - \Lambda'(\Phi^*)\big) U_c(i, \Phi^*) y(i, \Phi^*) \tag{11}$$

Thus, when $\tau'(\Phi^*) > \Lambda'(\Phi^*)$, the sensitivity of the surplus in marginal utility units across shocks, $\Omega_i'(\Phi^*)$, depends on the variation of output in marginal utility units, $U_c(i, \Phi^*)y(i, \Phi^*)$, at the balanced budget.

Part 1. Consider now the constant Frisch elasticity utility function. We showed in proposition **6** in the text that for $\rho > 1$, output in marginal utility units decreases as the shock increases, and therefore $\Omega_i'(\Phi^*)$ is decreasing in i. For $\rho < 1$, output in marginal utility units is procyclical, and therefore $\Omega_i'(\Phi^*)$ is increasing in i.

Expression (11) allows us to connect the monotonicity of $\Omega_i'(\Phi^*)$ to the IES, which holds for any number of shocks N. For the determination of distortions, we can connect the monotonicity of $\Omega_i'(\Phi^*)$ to $z_i'(\Phi^*)$ through lemma 3. We would like to show that if $\Omega_i'(\Phi^*)$ is increasing (decreasing) in i, then $z_i'(\Phi^*)$ is increasing (decreasing) in i. If the monotonicity of the sensitivity of surplus is bequeathed to the sensitivity of debt, we can use lemma 2 and talk about countercyclicality and procyclicality of distortions for the case of $\rho > 1$ and $\rho < 1$ respectively and get the results of the proposition. The result on the negative or positive drift under a worst-case model that assigns higher probability to bad (low TFP) shocks follows as in proposition **7**.

Given the monotonicity of $\Omega_i'(\Phi^*)$, the monotonicity of $z_i'(\Phi^*)$ depends in general on the persistence properties of the stochastic matrix $\mathbf{A} \equiv \Pi \circ \mathbf{M} \circ \mathbf{\Phi}$ in the present value formula (10). Let $N = 2$ and let the vector $\mathbf{y} = [y_1, y_2]^T$ be determined by the present value formula $\mathbf{y} = (I - \beta A)^{-1}\mathbf{x}$ with $\mathbf{x} = [x_1, x_2]^T$ and

$$\mathbf{A} \equiv \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}, \quad a, b \in (0,1).$$

We have then

$$\begin{aligned} y_1 &= \frac{1}{|I - \beta\mathbf{A}|}\big[(1-\beta b)x_1 + \beta(1-a)x_2\big] \\ y_2 &= \frac{1}{|I - \beta\mathbf{A}|}\big[\beta(1-b)x_1 + (1-\beta a)x_2\big] \end{aligned}$$

where $|I - \beta\mathbf{A}| = (1-\beta)\big[1 + \beta(1-(a+b))\big] > 0$ (the i.i.d. case corresponds to $a + b = 1$). This implies that $y_1 > y_2 \Leftrightarrow x_1 > x_2$. Reinterpret then $\mathbf{x}$ as $\mathbf{\Omega}'$ and $\mathbf{y}$ as $\mathbf{z}'$, and the result follows. Thus, if $\rho > 1$ we have $z_1'(\Phi^*) > z_2'(\Phi^*)$ and therefore $\Phi_{1|i}'(\Phi^*) > \Phi_{2|i}'(\Phi^*)$. Note that since $N = 2$ and since the derivatives $\Phi_{j|i}'(\Phi^*) > 0$ sum to unity by (6), we have $\Phi_{1|i}'(\Phi^*) > 1$ and $\Phi_{2|i}'(\Phi^*) < 1$. The opposite results hold for $\rho < 1$.

When we have more than two values of the shock, $N > 2$, z_i' will inherit the monotonicity of Ω_i' depending on the persistence properties of the matrix $\mathbf{A}$. It is obvious that if the induced measure (which is more complicated than the worst-case measure at the balanced budget since it depends on $\Phi_{j|i}'(\Phi^*)$) is i.i.d., then the monotonicity of $\Omega_i'(\Phi^*)$ is directly bequeathed to the present value of these coefficients, $z_i'(\Phi^*)$, for any $N > 2$. The same holds if the induced measure is also very persistent (which is something we expect).[2]

Part 2. The entire expansion is valid for any kind of period utility function that generates a tax rate and a government share that are functions solely of Φ, i.e. $\tau_t = \tau(\Phi_t), \Lambda_t = \Lambda(\Phi)$. In that case, Φ^* such that $\tau(\Phi^*) = \Lambda(\Phi^*)$ is always a fixed point of the law motion and all the results up to now can be used.

Consider now the utility function $U = \frac{u^{1-\rho}-1}{1-\rho}$, where $u = c^{\alpha_1} l^{\alpha_2} g^{\alpha_3}, \quad \alpha_i > 0, \sum_i \alpha_i = 1$, which satisfies balanced growth restrictions for the case also for $\rho \neq 1$. We show first that τ and Λ are only functions of Φ. For these preferences the intratemporal marginal rates of substitution take the form $\frac{U_l}{U_c} = \frac{\alpha_2}{\alpha_1}\frac{c}{l}$ and $\frac{U_g}{U_c} = \frac{\alpha_3}{\alpha_1}\kappa^{-1}, \kappa \equiv g/c$. The elasticities of the utility function are

$$\begin{aligned}
\epsilon_{cc} &= 1 - \alpha_1(1-\rho) \\
\epsilon_{ch} &= \alpha_2(1-\rho)\frac{h}{l} \\
\epsilon_{hh} &= (1 - \alpha_2(1-\rho))\frac{h}{l} \\
\epsilon_{hc} &= \alpha_1(1-\rho) \\
\epsilon_{gc} &= \alpha_1(1-\rho) \\
\epsilon_{gh} &= -\alpha_2(1-\rho)\frac{h}{l}
\end{aligned}$$

Remember that the public wedge for $t \geq 1$ depends on the elasticities as follows:

$$\chi = \frac{\Phi(1 - \epsilon_{cc} - \epsilon_{ch} - \epsilon_{gc} - \epsilon_{gh})}{1 + \Phi(\epsilon_{gc} + \epsilon_{gh})}$$

and note that

$$\epsilon_{cc} + \epsilon_{ch} + \epsilon_{gc} + \epsilon_{gh} = 1.$$

Thus, $\chi = 0$ and the government share is the same as in the first-best. In particular, we have $\kappa^{FB} = \frac{\alpha_3}{\alpha_1}$ and $\Lambda = \Lambda^{FB} \equiv \frac{\alpha_3}{\alpha_1 + \alpha_3}$. So we have a constant Λ independent of Φ.[3]

The optimal tax rate depends on the labor/leisure ratio as follows:

[2] In numerical experiments we played around with $N = 11$ and we always faced the case where the monotonicity of Ω_i' was bequeathed. The induced matrix $\mathbf{A}$ was always very persistent.

[3] We were getting the same result for the basic parametric example of the paper when we had unitary elasticity of substitution between c and g, $\psi = 1$. So the zero public wedge result extends for the non-separable case when we allow also unitary elasticity substitution with leisure.

$$\tau = \frac{\Phi(1+\frac{h}{l})}{1+\Phi(1+\frac{h}{l}+(1-\rho)[\alpha_1-\alpha_2\frac{h}{l}])}$$

Since we do not have a constant Frisch elasticity as in the basic parametric example, the tax rate could in principle depend through labor on the shock s. This is not the case though. To see that, consider the optimal wedge equation that takes the form

$$\frac{U_l}{U_c} = \frac{\alpha_2}{\alpha_1}\frac{c}{l} = \frac{1+\Phi(1-\epsilon_{cc}-\epsilon_{ch})}{1+\Phi(1+\epsilon_{hh}+\epsilon_{hc})}s$$

But since $c = (1-\Lambda)y = (1-\Lambda)sh$, we can eliminate the technology shock and finally get the following equation

$$\frac{\alpha_2}{\alpha_1}(1-\Lambda)\frac{h}{l} = \frac{1+\Phi(1-\rho)(\alpha_1-\alpha_2\frac{h}{l})}{1+\Phi(1+\frac{h}{l}+(1-\rho)[\alpha_1-\alpha_2\frac{h}{l}])}$$

This equation determines a quadratic in h/l which allows to solve for labor as a function of Φ, $h = h(\Phi)$. Thus, the optimal tax rate becomes function only of Φ, $\tau(\Phi)$, and not of the shock s. The reason behind this result is obviously the fact that the income and substitution effect in labor supply cancel out for these preferences, making labor constant (given a constant Φ). Note that output is then $y = sh(\Phi)$ and that the surplus takes the form $S = (\tau(\Phi)-\Lambda)y(s,\Phi)$. The balanced budget Φ^* satisfies $\tau(\Phi^*) = \Lambda = \Lambda^{FB}$. Thus, the balanced budget approximation can be used.

To finish the proof of part **2**, we need to associate the IES of the composite good $1/\rho$ to the allocation of distortions. Note that marginal utility takes the form $U_c = \alpha_1 c^{\alpha_1(1-\rho)-1} l^{\alpha_2(1-\rho)} g^{\alpha_3(1-\rho)}$. Using $c = (1-\Lambda)y, g = \Lambda y$ and the fact that leisure is only function of Φ, this can be rewritten as

$$\begin{aligned} U_c &= K(\Phi)\cdot y^{(\alpha_1+\alpha_3)(1-\rho)-1} \\ K(\Phi) &\equiv \alpha_1(1-\Lambda)^{\alpha_1(1-\rho)-1}\Lambda^{\alpha_3(1-\rho)} l(\Phi)^{\alpha_2(1-\rho)} > 0 \end{aligned}$$

Thus, the optimal surplus in marginal utility units as function of the shock i and the excess burden of taxation Φ is

$$\Omega_i(\Phi) = K(\Phi)(\tau(\Phi)-\Lambda)y(i,\Phi)^{(\alpha_1+\alpha_3)(1-\rho)}$$

Thus, for $\tau(\Phi) > \Lambda$, the surplus in marginal utility units is procyclical when $\rho < 1$ and countercyclical when $\rho > 1$, so the results of proposition **6** go through. Furthermore, since $\tau'(\Phi^*) > \Lambda'(\Phi^*) = 0$, the monotonicity of $\Omega_i'(\Phi^*)$ depends on output in marginal utility units, as seen from expression (11) (there was no assumption for the utility function for its derivation). Given our derivation above, the sensitivity with respect to shock s is controlled again by the parameter ρ: $\Omega_i'(\Phi^*)$ is increasing in i if $\rho < 1$ and decreasing in i if $\rho > 1$. The results of part **1** follow.

Period utility at the balanced budget. The results about the drifts according to the reference measure in propositions **7** and **8** are based on the assumption (which always holds numerically) that at the balanced budget the worst-case measure assigns higher probability on low technology shocks. This would be so if we could show that $V_i(\Phi^*)$ is increasing in shock i in (1). We will show here that the period utility function is an increasing function of the shock, so $U_i(\Phi^*)$ is increasing in i. We show this result for any kind of utility functions that generate optimally a τ and Λ that are solely functions of Φ, so both of our parametric examples are covered.

$$\mathcal{V}(s,\Phi) \equiv U(c(s,\Phi), 1-h(s,\Phi), g(s,\Phi)) = U\big((1-\Lambda(\Phi))y(s,\Phi), 1-h(s,\Phi), \Lambda(\Phi)y(s,\Phi)\big).$$

We obviously have $U_i(\Phi^*) = \mathcal{V}(s_i, \Phi^*)$. Differentiate with respect to s to get

$$\begin{aligned}\frac{\partial \mathcal{V}}{\partial s} &= U_c(1-\Lambda(\Phi))\frac{\partial y}{\partial s} - U_l\frac{\partial h}{\partial s} + U_g\Lambda(\Phi)\frac{\partial y}{\partial s}\\ &= U_c\Big[\frac{\partial y}{\partial s} - \frac{U_l}{U_c}\frac{\partial h}{\partial s} + \Lambda(\Phi)[\frac{U_g}{U_c} - 1]\frac{\partial y}{\partial s}\Big]\end{aligned}$$

Use now $U_l/U_c = (1-\tau)s$ and $U_g/U_c = 1+\chi$ to get

$$\frac{\partial \mathcal{V}}{\partial s} = U_c\Big[\frac{\partial y}{\partial s} - (1-\tau)s\frac{\partial h}{\partial s} + \Lambda(\Phi)\chi\frac{\partial y}{\partial s}\Big]$$

Now, note that $\partial y/\partial s = h + s\partial h/\partial s$. Use this fact to get

$$\frac{\partial \mathcal{V}}{\partial s} = U_c\big[(\tau(\Phi) + \chi\Lambda(\Phi))\frac{\partial y}{\partial s} + (1-\tau(\Phi))h\big]$$

Note that there could be a potentially negative effect of s to the period utility if there is a negative public wedge (or a labor subsidy – which is not optimal for our parametric examples). At the balanced budget the expression simplifies to

$$\frac{\partial \mathcal{V}^*}{\partial s} = U_c^* \left[(\tau(\Phi^*)(1+\chi^*)\frac{\partial y^*}{\partial s} + (1-\tau(\Phi^*))h^* \right] > 0,$$

since $1+\chi = U_g/U_c > 0$ and $\partial y/\partial s > 0$. Thus, $U_i(\Phi^*)$ is increasing in i.

A.6 An algorithm

The approximation can be used also for computational purposes, as long as we stay in the vicinity of the balanced budget. To see how, we sketch here an algorithm.

Solve first for the worst-case measure at the balanced budget $m^*_{j|i}$, by calculating utilities from recursion (1). Solve afterwards for N^2+N unknowns ($\Phi'_{j|i}(\Phi^*)$ and $z'_i(\Phi^*)$) from N^2+N equations ((3) and (10)) through the following iterative procedure:

- Make a guess for $\mathbf{\Phi}$. Derive induced derivatives of the relative debt positions $\mathbf{z}'$ from (10).
- Use $\mathbf{z}'$ to get the Jacobian $J^*_{\eta_i}, \forall i$ from (4) and update the guess for $\mathbf{\Phi}$ by solving the systems (3).
- Iterate till convergence.

We use as a first guess $\mathbf{\Phi}_0 = \mathbf{1}_{N\times N}$. When updating the guess we also use damping in order to improve the convergence properties of the loop. For small σ (in absolute value), we could find a solution that was also robust to different initial guesses. For large σ though the non-convexities of the problem become pronounced and there is no guarantee of convergence of the algorithm. We used this algorithm for $N = 11$ and for the various calibrations used in the text. Results are available among request. The text features results from the global solution for $N = 2$.

It is sufficient to use the linear approximation around Φ^* only for the excess burden of taxation and for the debt in marginal utility units z. We choose the initial value Φ_0 and the initial allocation (c_0, h_0, g_0) by using the optimality conditions at $t = 0$ and requiring that the implementability constraint at $t = 0$ holds. For the allocation and policy at $t \geq 1$, we can use the non-linear functions for $(\tau(\Phi), \Lambda(\Phi))$ and $(c(\Phi), h(\Phi), g(\Phi))$, where Φ follows the approximate law of motion (2). So the method we illustrate is "hybrid".

B Government consumption share

In our baseline experiments we abstracted from variation in the government consumption share Λ and focused on $\psi = 1$. Consider now the case of substitutes ($\psi < 1$) and complements ($\psi > 1$). We consider four pairs of (ρ, ψ) and calibrate all other parameters as previously. For each pair, we always re-calibrate (α, a_h), so that the same first-best government share and labor are targeted.

Table 1: Correlation of $\Delta\Lambda$ with the technology shock.

	Substitutes ($\psi = 0.9$)	Complements ($\psi = 1.1$)
Low IES ($\rho = 2$)	0.4884	-0.5364
High IES ($\rho = 0.5$)	-0.5883	0.5543

The table depicts $\mathrm{Corr}(\Delta\Lambda, s)$ for 4 different sets of (ρ, ψ). For each set of parameters we generated 10,000 sample paths of 200-period length. The reported numbers are mean statistics across sample paths.

Table 1 displays the correlations of Λ with technology shocks. Recall from our analysis in proposition **3** that a higher distortion (in the sense of Φ) implies a lower (higher) government share Λ when we have substitutes (complements). Consider first the case of a low IES ($\rho > 1$), where distortions are *negatively* correlated with the cycle and exhibit a negative drift. High distortions in bad times and low distortions in good times imply a government share that decreases in bad times and increases in good times if we have substitutes. The opposite happens for the complements case.

So, changes in Λ are procyclical (countercyclical) if we have $\psi < 1$ ($\psi > 1$), as the first row of table 1 shows. Furthermore, since the excess burden is reduced on average over time till its balanced-budget value Φ^* is reached, the respective distortions at the provision of government consumption are also reduced till the rest point $\Lambda(\Phi^*)$. Hence, in the case of substitutes, where Λ is initially *below* its balanced-budget value, we have a positive drift of the government share over time. Consequently, front-loaded taxes are accompanied with back-loaded government expenditures. In contrast, in the case of complements, where the share of government consumption is initially above its balanced-budget value, Λ exhibits a negative drift over time.

When the IES is high ($\rho < 1$), distortions are procyclical and exhibit a positive drift over time. Obviously, a higher distortion when the technology shock is high implies then a lower Λ in the substitutes case and a higher Λ in the complements case, which explains the sign of the correlations in the second row of the table. Similarly, Λ exhibits a negative drift for $\psi < 1$ and a positive drift when $\psi > 1$.

Figure 1 summarizes the mean dynamics of the government share. We note that the changes in the government share over time are small for all pairs of (ρ, ψ), a fact which may justify the focus on $\psi = 1$.

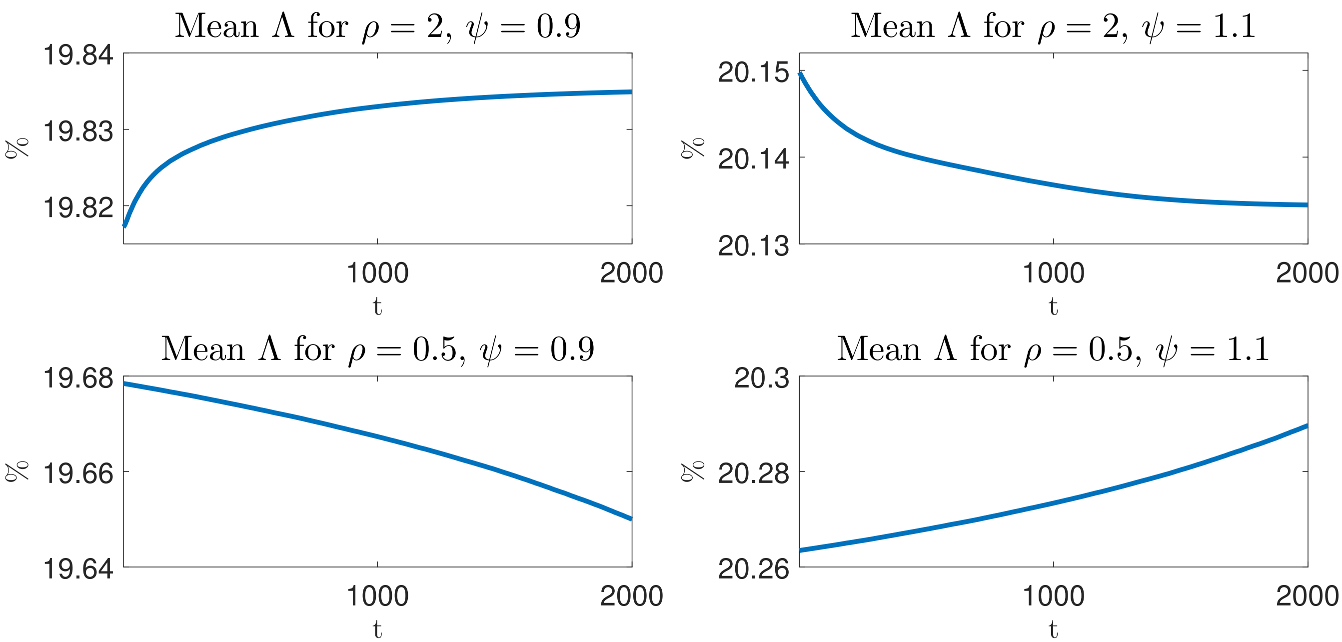

Figure 1: Evolution of the mean government share over time. The two graphs on the top consider the case of $\rho = 2$. The two graphs on the bottom consider the case of $\rho = 0.5$. Graphs on the left correspond to the substitutes case ($\psi = 0.9$) and graphs on the right to the complements case, ($\psi = 1.1$). When $\rho = 2$ we have convergence to the balanced-budget government share that is either below (substitutes) or above (complements) the first-best government share of 20%. When $\rho = 0.5$, the government share diverges.