FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Fiscal Austerity in Ambiguous Times

Axelle Ferrière and Anastasios G. Karantounias

Working Paper 2016-6
March 2016

Abstract: How should public debt be managed when uncertainty about the business cycle is widespread and debt levels are high, as in the aftermath of the last financial crisis? 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 optimal policies that resemble "austerity" measures. Optimal policy prescribes front-loaded fiscal consolidations and convergence 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: endogenous government expenditures, distortionary taxes, balanced budget, austerity, fiscal consolidation, martingale, ambiguity aversion, multiplier preferences

---

The authors are grateful to David Backus, Richard Rogerson, Thomas J. Sargent, Stanley E. Zin, and especially to Mikhail Golosov for their comments. They also thank Anmol Bhandari, Murat Tasci, conference participants at the Society for Economic Dynamics meetings in Toronto, the European Economic Association meetings in Toulouse, the Computing in Economics and Finance meetings in Oslo, the Ambiguity and Robustness workshops at the Becker Friedman Institute and at New York University, and seminar participants at the Federal Reserve Banks of Atlanta and Cleveland. 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, 43 via della Piazzuola, 50133, Firenze, 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, anastasios.karantounias@atl.frb.org.

1 Introduction

Historically-high levels of public debt have triggered heated debates regarding debt sustainability and the appropriateness of austerity measures in the aftermath of the financial crisis of 2007-2009. In particular, is it optimal to reduce debt? And if so, at which pace? Also, what should be the main instrument of debt reduction, higher distortionary taxes or lower government spending? Satisfying answers are not easy to come up with, especially in the context of uncertainty about the strength of economic recovery. To shed light on these questions, we analyze the optimal determination of government spending, taxes and debt in an environment of *ambiguity* about the cycle. Our main finding is that *austerity* can become optimal in such an economy if interest rates are sufficiently responsive to cyclical shocks. Optimal policy prescribes then front-loaded fiscal consolidations and convergence to a balanced primary budget in the long-run.

Our environment features an economy without capital and complete markets as in Lucas and Stokey (1983). We endogenize though government consumption by allowing it to provide utility to the representative household. Our government faces an initial stock of debt and uses distortionary labor taxes and state-contingent debt for its financing needs. 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 provide a comprehensive analysis of optimal policy in the case of an expected-utility household that has full confidence in the probability model. We define a new wedge at the second-best, which we call the *public wedge*. This wedge captures the deviation of the marginal rate of substitution of government for private consumption from unity. The public wedge can be either positive or negative, reflecting respectively government expenditures that may be low or high relative to the first-best.

We derive several lessons from the analysis of optimal dynamic policy without ambiguity. Our first finding is that the optimal allocation, public wedge and tax rate are history-independent, reflecting how the smoothing motives in the Lucas and Stokey setup extend to environments where government expenditures provide utility. Using a homothetic specification for the utility of private and government consumption results in a share of government consumption in output that is not only history-independent, but actually constant. Hence, optimal government expenditures are *procyclical* whereas their share becomes *acyclical*. Furthermore, the tax rate is constant under the assumption of a constant Frisch elasticity of labor supply. Optimal policy prescribes a deficit at the initial period and then a constant surplus-to-output ratio afterwards. Public debt remains stationary, without exhibiting negative or positive drifts. Consequently, *neither* fiscal consolidations *nor* additional accumulation of public debt are optimal.

There are stark differences when we turn to the analysis of the optimal fiscal policy in an environment of 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 sensitivity of interest rates to consumption growth. This sensitivity is controlled by the intertemporal elasticity of substitution (IES). In particular, 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 tax rates till debt becomes zero and a balanced primary budget is reached in the long-run. These two facets of optimal policy constitute the "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 of 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 taxes, by reducing the utility of the household, raise the pessimistic probabilities and therefore *increase* equilibrium prices of state-contingent claims. This return-reducing effect of taxes is used by the planner to *amplify* the present value of the *portfolio* of newly issued government securities, a policy that relaxes the government budget constraint and increases welfare.

Instrumental in the calculation of present values is the behavior of surpluses, and consequently of debt, in *marginal utility* units. Surpluses may fall in bad times, but contractions of output are accompanied by expansions in marginal utility and therefore a decrease in state-contingent returns. If the IES is lower than unity, then marginal utility and effectively interest rates are sufficiently responsive to shocks, leading to surpluses and debt in marginal utility units that are actually high in recessions and low in expansions. Since the present value of debt consists of the product of the pessimistic beliefs and debt in marginal utility units, high taxes in bad times and low taxes in good times, amplify, through the channel of pessimistic beliefs, the value of the high debt positions contingent on bad times and reduces the value of the low debt positions contingent on good times, increasing therefore the overall value of the government portfolio. The opposite happens when the IES is larger than unity: marginal utility is not responsive enough, leading to debt in marginal utility units that is procyclical. A procyclical tax rate then increases the value of the government portfolio. In the knife-edge case of a unitary IES, debt in marginal utility units is constant across the cycle, muting therefore the desire of the planner to increase the market value of debt by managing the pessimistic expectations of the household and leading to the same fiscal

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

policies as without ambiguity.

The long-run results of a negative drift in public debt till a balanced budget is reached, or of further accumulation of public debt are based on the assumption that doubts about the model are unfounded, i.e. the probability model that the agents doubt is actually the true data-generating process. We show that higher doubts about the model, or higher persistence and volatility of technology shocks lead to more aggressive expectation management and sharper fiscal consolidations when the IES is smaller than unity.

The intuition behind the drifts in taxes and debt relies on the difference between the pessimistic beliefs and the actual probability of shocks. Good times bear low taxes when the IES is below unity. 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 of 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.

The behavior of the government share in output is more nuanced in an environment with ambiguity, because it depends on the substitutability of government with private consumption. Higher distortions at the government consumption margin may imply a government share in output that falls (in the case of substitutes) or rises (in the case of complements). Consequently, when the IES is below unity (our "austerity" case) and therefore distortions increase in bad times, we have a government share that falls (rises) in bad times in the case of substitutes (complements). Furthermore, the government share converges to its balanced-budget value starting from a relatively low (high) share for the case of substitutes (complements). For the high IES case, the opposite picture emerges: the government share falls (rises) in good times in the case of substitutes (complements). Moreover, it becomes either progressively smaller (substitutes) or progressively larger (complements) over time. So if we adopt the view that government and private consumption are substitutes, then the share of government consumption is optimally procyclical in the austerity case and countercyclical for the opposite, anti-austerity case.

1.1 Related literature

Optimal taxation studies typically treat government expenditures as exogenous, abstracting from –relevant for fiscal consolidations– questions about the optimal mix of taxes and spending.[2] Bachmann and Bai (2013) perform a positive analysis 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 in answering questions about public debt.[3]

[2]Studies like Teles (2011) raise concerns about this practice.

[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.

There are various studies 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 inequality and lack of commitment can *optimally* lead to cycles between austerity and populistic regimes.

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 (2015) on sovereign default, whereas Croce et al. (2012) studied the effects of technological and fiscal uncertainty on long-run growth.

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 can lead also to interesting inertia as in the work of Ilut et al. (2016).[4]

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 where 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. 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. Furthermore, both the prominent role of the IES and the fiscal *austerity* result do not feature in Karantounias (2013b). The deep reason behind these differences stems from the fact that by endogenizing

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

government consumption and adopting a homothetic specification, we get a clean dependence of the equilibrium on the IES *only* and not on features of the economy that break homotheticity like the size of exogenous government expenditures.

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 exhibits the problem of a planner that adopts the welfare criterion of the household. Section 5 analyzes the allocation of distortions over states and dates and section 6 proceeds to the numerical evaluation of optimal policy. Section 7 concludes. The Appendix provides proofs of the propositions and details about the particular expansion we used to solve the policy problem 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. We will explore specifications of U later. 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 $q_0 \equiv 1$.

There is a representative competitive firm that operates the linear technology. The government is collecting tax revenues to finance government consumption 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.[5]

[5]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 $b_t^g = -b_t$. Using the resource constraint and the household's budget constraint delivers in equilibrium the government budget constraint, so it is redundant to include it in the definition of the competitive equilibrium.

2.1 Household's optimality conditions

Profit maximization of the competitive firm leads to a wage that is equal to the marginal product of labor, $w_t = s_t$. Turning to the household's problem, its labor supply decision is characterized by

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

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

The household takes government expenditures as exogenously given in the competitive equilibrium. Consider now the problem of the Ramsey planner that chooses 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.

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, at the first-best, the optimal provision of government consumption requires that it provides the same marginal utility as private consumption. Equation (9) is standard; it 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) by expressing prices and tax rates in terms of allocations, and have the Ramsey planner choose (c, h, g) subject to the resource constraint and the associated implementability constraints that guarantee that the second-best allocation can be supported by a competitive equilibrium. Define first

$$\Omega(c, h, g) \equiv U_c(c, 1-h, g)c - U_l(c, 1-h, g)h. \tag{10}$$

The function Ω stands for consumption net of after-tax labor income in marginal utility of consumption units, after expressing the after-tax wage in terms of allocations through (6). Note that Ω is also equal to the primary government surplus in marginal utility units. Using the household's (or equivalently the government's) intertemporal budget constraint and substituting intertemporal marginal rates of substitution for equilibrium prices and intratemporal marginal rates of substitution for after-tax wages delivers the familiar implementability constraint

$$\sum_{t=0}^{\infty} \beta^t \sum_{s^t} \pi_t(s^t) \Omega(c_t(s^t), h_t(s^t), g_t(s^t)) = U_c(c_0, 1-h_0, g_0) b_0. \tag{11}$$

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 (11) and the resource constraint (1), where the initial shock s_0 and initial debt b_0 are given.*

Let Φ denote the multiplier on the unique implementability constraint. We call Φ the excess burden of taxation throughout the paper. The first-order conditions of the second-best problem are stated in the Appendix. They imply the following two expressions for the government consumption margin and the labor supply margin for $t \geq 1$,

$$\frac{U_g + \Phi\Omega_g}{U_c + \Phi\Omega_c} = 1 \tag{12}$$

$$\frac{U_l - \Phi\Omega_h}{U_c + \Phi\Omega_c} = s, \tag{13}$$

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

Expressions (12) and (13) 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 $\Phi = 0$).[6] Before we proceed to an analysis of the wedges, it is useful to note that using (12) and (13) 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, $c_t = c(s_t, \Phi), h_t = h(s_t, \Phi), g_t = g(s_t, \Phi), t \geq 1$. Thus,

Proposition 1. *The optimal allocation (c, h, g) is history-independent.*

This proposition extends the basic result of Lucas and Stokey (1983) to environments with endogenous government consumption.

3.3 Optimal public wedge and labor tax

Define

$$\chi \equiv \frac{U_g}{U_c} - 1. \tag{14}$$

We will 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, which is unity. In particular, if $\chi > 0$, the marginal utility of government consumption is larger than the marginal utility of private consumption, $U_g/U_c > 1$, which we will interpret as having a small government consumption relative to the first-best. Similarly, we will interpret a negative public wedge $\chi < 0$ as a situation where government consumption is large relative to the first-best.

Note that 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, $\chi_t = \chi(s_t, \Phi), \tau_t = \tau(s_t, \Phi)$. It is useful to express the optimal χ and τ as functions of *elasticities* that capture curvature properties

[6] The respective wedges at $t = 0$ are stated in the Appendix.

of the period utility function U, and the excess burden of taxation Φ.[7]

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 curvature properties of the utility function show up in the determination of the wedges because they capture how the surplus in marginal utility units Ω –the main ingredient in the calculation of present values– is affected by the choices of c, h and g. The proposition shows 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.

3.4 Parametric example

Consider the period utility function

$$U = \frac{u^{1-\rho} - 1}{1 - \rho} + v(l), \tag{15}$$

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

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 specification of the period utility function 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 understanding the properties of the optimal plan under ambiguity, as will become clear later. We will call c and g *substitutes* when $\psi < 1$ and *complements* when $\psi > 1$.[8] In the same vein, we are going to talk about *intertemporal* substitutability in terms of composite consumption when $\rho < 1$ and intertemporal complementarity when $\rho > 1$.

The homothetic specification in private and government consumption allows us to perform our analysis in terms of *ratios*. The marginal rate of substitution of government for private consumption is $U_g/U_c = A(g/c)^{-\psi}$, where $A \equiv \alpha/(1-\alpha)$. Let κ denote the ratio of government to private consumption, $\kappa \equiv g/c$, and let $\Lambda \equiv g/y = \kappa/(1+\kappa)$ denote the share of government consumption in output. At the first-best we have we have $\kappa = \kappa^{FB} \equiv A^{1/\psi}$ with a government share $\Lambda^{FB} \equiv \kappa^{FB}/(1+\kappa^{FB})$. At the second-best, we have $\kappa < \kappa^{FB}$ and a share of government consumption that is small relative to the first-best, $\Lambda < \Lambda^{FB}$, when there is a positive public wedge ($\chi > 0$). In the case of a negative public wedge ($\chi < 0$), we get $\kappa > \kappa^{FB}$ and a share of government consumption that is large relative to the first-best, $\Lambda > \Lambda^{FB}$.

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 $\lambda_c \equiv (1-\alpha)(\frac{c}{u})^{1-\psi} \in (0,1)$.[9] Therefore, $\epsilon_{cc} + \epsilon_{gc} = \psi$, so the public wedge in proposition 2 becomes

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

As stated in proposition 2, the sign of χ, and therefore, the size of Λ relative to the first-best,

[8] 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).

[9] 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 $\psi = 1$.

is determined by the numerator. It is important to note that the optimal Λ is not necessarily small relative to the first-best, despite the fact that government consumption has to be financed through distortionary taxes. For the Cobb-Douglas case of $\psi = 1$, the planner does *not* distort the government consumption margin and sets a zero public wedge (implying the first-best ratios $\kappa^{FB} = A, \Lambda^{FB} = \alpha$; *levels* of g are of course different from the first-best). In contrast, there is a positive public wedge and a small share Λ relative to the first-best, in the case of *substitutes* ($\psi < 1$), and a negative public wedge and a large share Λ relative to the first-best in the case of *complements* ($\psi > 1$). Thus, the intratemporal elasticity of substitution is instrumental in the determination of distortions in the provision of government consumption.

The following proposition collects further results for the second-best allocation and the induced tax and government consumption policy for $t \geq 1$.

Proposition 3. *1. Assume the homothetic specification in (15). Then,*

(a) The ratio κ is function only of Φ and not of the shocks s. Thus, the share of government consumption in output Λ is function only of Φ, $\Lambda(\Phi)$, and, therefore, <u>constant</u> 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 that $v(l) = -a_h \frac{(1-l)^{1+\phi_h}}{1+\phi_h} = -a_h \frac{h^{1+\phi_h}}{1+\phi_h}$. Then,

(a) The tax rate is function only of Φ, $\tau(\Phi)$, and therefore is <u>constant</u> across shocks.

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

(c) Equilibrium labor and output are $h(s, \Phi) = H(\Phi) \cdot s^{\frac{1-\rho}{\rho+\phi_h}}$ and $y(s, \Phi) = H(\Phi) \cdot s^{\frac{1+\phi_h}{\rho+\phi_h}}$, where $H(.)$ is defined in the Appendix.

(d) When $\tau > \Lambda$, the surplus is increasing in s.

(e) Let $\Omega^\star(s, \Phi)$ denote the optimal surplus in marginal utility units as a function of (s, Φ). For the current utility specification it takes the form

$$\Omega^\star(s, \Phi) = \big(\tau(\Phi) - \Lambda(\Phi)\big) J(\Phi) \cdot \big[y(s, \Phi)\big]^{1-\rho}, \tag{16}$$

where $J(.)$ is defined in the Appendix and $J(\Phi) > 0$. Thus, when $\tau > \Lambda$, we have $\operatorname{sign} \partial\Omega^\star / \partial s = \operatorname{sign}(1 - \rho)$.

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 perfect tax-smoothing result arises. The proposition shows formally how increases in the excess burden of taxation Φ correspond to larger distortions at the government consumption and the labor supply margin, for either particular parameter configurations, or, more generally, for small deviations from the first-best. 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.

Note that the constant share of government consumption and the constant tax rate lead to a neat multiplicative separability of labor in s and Φ. Income and substitution effects in labor supply are controlled only by ρ, although the marginal utility of consumption is controlled by both ρ and ψ. Regarding the properties of primary surpluses we have the following result: surpluses are procyclical whereas the surplus-to-output ratio is *acyclical*. It is crucial for later purposes to understand the behavior of surpluses in marginal utility units $\Omega^\star$ as function of the technology shocks. The expansion in output due to a positive technology shock and therefore the increase in surplus is counteracted by the contraction of marginal utility due to an increase in consumption. The decrease is controlled by ρ, the inverse of IES. If $\rho > 1$, so if $IES < 1$, we have surpluses in marginal utility units that are countercyclical. In contrast, when $\rho < 1$ and therefore $IES > 1$, Ω^* is procyclical. Note that at the knife-edge case of $\rho = 1$, $\Omega^\star$ is constant across shocks. It is also easy to see that a zero initial debt leads to a balanced budget for $t \geq 0$.

Proposition 4. *("Optimality of balanced budgets"). Let the utility function be as in (15) 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. □

Illustration. Figure 1 plots the optimal tax rate and the government share Λ for the substitutes and the complements case as functions of Φ. Our calibration is standard. We use an annual frequency and a unitary Frisch elasticity, so we set $(\beta, \phi_h) = (0.96, 1)$. The size of ρ in the full confidence case does not play any role in the dynamics of τ and Λ and therefore we set it equal to unity, $\rho = 1$. 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 approximate the AR(1) process with 11 points using the Rouwenhorst method of Kopecky and Suen (2010). For each ψ we set α and a_h so that at the first-best the government share is 20% and the household works 40% of its time. These values correspond to equilibrium labor that is about 35.5% of the available time at the second-best. The initial shock is unity, $s_0 = 1$, and

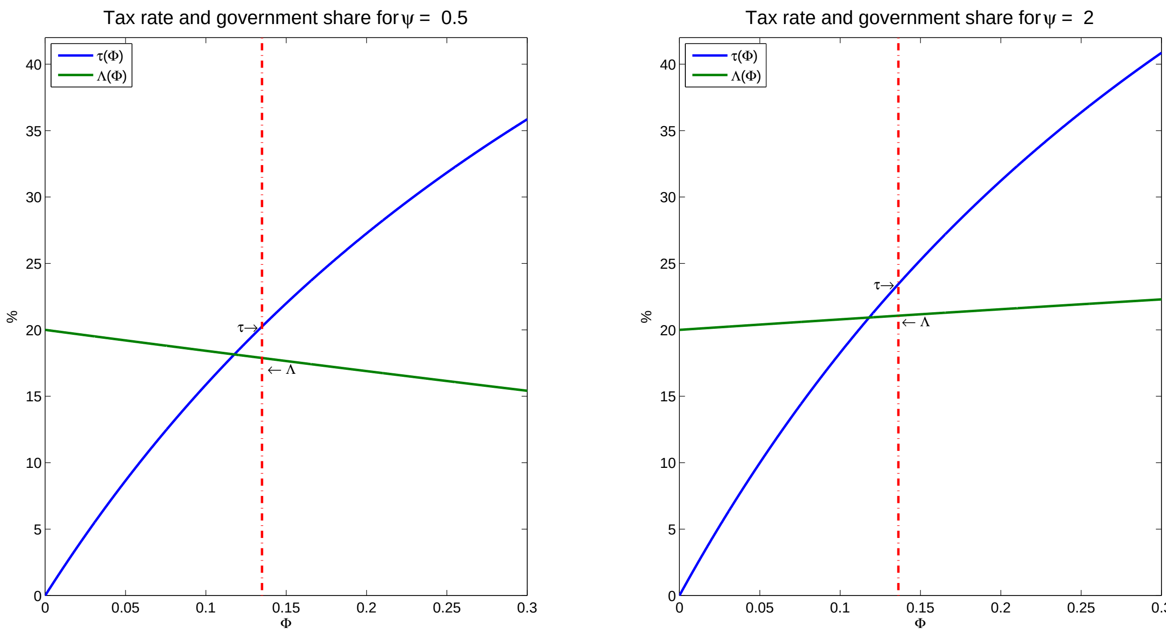

Figure 1: Plots of $\tau(\Phi)$ and $\Lambda(\Phi)$ for the substitutes (left graph) and the complements case (right graph). The vertical line indicates the Φ that satisfies the intertemporal budget constraint with positive initial debt and denotes the optimal tax rate and Λ for $t \geq 1$. The intersection of τ and Λ depicts the balanced budget policy.

Table 1: Optimal policy for the substitutes and complements case.

	$\psi = 0.5$	$\psi = 2$		$\psi = 0.5$	$\psi = 2$
	$\mathbf{t = 0}$			$\mathbf{t \geq 1}$	
τ_0	13.97	14.87	τ	20.25	23.45
Λ_0	16.77	19.64	Λ	17.88	21.07
b_0/y_0	54.42	54.44	b/y	59.24	59.53

All variables are expressed in %. The constancy of the debt-to-output ratio is a special outcome of the $\rho = 1$ assumption. More generally, it would depend on the current chock s.

initial debt is $b_0 = 0.2$, which corresponds to 50% of first-best output. The figure shows that the tax rate is an increasing function of the excess burden of taxation whereas the government share is decreasing for the substitutes case ($\psi = 0.5$) and increasing for the complements case ($\psi = 2$), as we expect from proposition 3. The balanced budget policy of proposition 4 is at the intersection of the τ-schedule and Λ-schedule. With positive initial debt, the government runs a deficit at $t = 0$ and then a constant surplus-to-output ratio for each $t \geq 1$. Table 1 reports the exact tax rates, government shares and debt ratios.

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 positive 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,[10]

$$V_t = U(c_t, 1-h_t, g_t) + \beta \min_{m_{t+1} \geq 0, E_t m_{t+1}=1} [E_t m_{t+1} V_{t+1} + \theta E_t m_{t+1} \ln m_{t+1}], \tag{17}$$

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 expected 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 (17) delivers the worst-case conditional change of measure

$$m_{t+1}(s^{t+1}) = \frac{\exp(\sigma V_{t+1}(s^{t+1}))}{\sum_{s_{t+1}} \pi_{t+1}(s^{t+1}|s^t) \exp(\sigma V_{t+1}(s^{t+1}))} \tag{18}$$

where $\sigma \equiv -\theta^{-1} < 0$, with $\sigma = 0$ corresponding to the expected utility case. Expression (18) shows that a cautious 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 (18) in (17) delivers the familiar risk-sensitive recursion

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

$$V_t = U(c_t, 1 - h_t, g_t) + \frac{\beta}{\sigma} \ln E_t \exp(\sigma V_{t+1}). \tag{19}$$

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

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

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 the continuation utility of the household and therefore, through the household's endogenous beliefs, equilibrium prices. In turn, equilibrium prices determine the desirability of debt and therefore 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.[11] We follow a recursive representation 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:

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

subject to

[11]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 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).

$$z = \Omega(c,h,g) + \beta \sum_{s'} \pi(s'|s) \frac{\exp(\sigma V(z'_{s'}, s'))}{\sum_{s'} \pi(s'|s) \exp\left(\sigma V(z'_{s'}, s')\right)} z'_{s'} \tag{20}$$

$$c + g = sh \tag{21}$$

$$c, g \geq 0, h \in [0,1], z'_{s'} \in Z(s') \tag{22}$$

Let Φ denote the multiplier on the dynamic implementability constraint (20) and let λ denote the multiplier on the resource constraint (21). The first-order necessary conditions are

$$c: \qquad U_c + \Phi\Omega_c = \lambda \tag{23}$$

$$h: \qquad -U_l + \Phi\Omega_h = -\lambda s \tag{24}$$

$$g: \qquad U_g + \Phi\Omega_g = \lambda \tag{25}$$

$$z'_{s'}: \qquad V_z(z'_{s'}, s')[1 + \sigma\eta'_{s'}\Phi] + \Phi = 0 \tag{26}$$

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, $m'_{s'} = \frac{\exp(\sigma V(z'_{s'}, s'))}{\sum_{s'} \pi(s'|s) \exp(\sigma V(z'_{s'}, s'))}$.

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 $\sum_{s'} \pi(s'|s) m'_{s'} = 1$.

4.4 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 (26) in sequence notation as

$$\frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} + \sigma\eta_{t+1}, t \geq 0 \tag{27}$$

where $\eta_{t+1} = z_{t+1} - E_t m_{t+1} z_{t+1}$.

The excess burden of taxation moves according to the relative debt position η_{t+1}, a feature that we will analyze in detail in the next section. Using the optimality conditions of the recursive problem (23), (24) and (25) delivers the two equations that characterize the optimal wedges at the two margins, (12) and (13), with the crucial difference that the excess burden of taxation is indexed by time, Φ_t.[12] This fact allows us to solve for (c, h, g) in terms of the shock s_t and a time-varying Φ_t, $c_t = c(s_t, \Phi_t), h_t = h(s_t, \Phi_t), g_t = g(s_t, \Phi_t)$. These functions are exactly the same functions of (s, Φ) as in the case without doubts about the model. Similarly, we have:

Proposition 5. *(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 (27).*

As a result, the formulas that we derived for the parametric example in proposition 3 go through by replacing the constant Φ with a time-varying Φ_t. An immediate implication is that the share of government consumption, the public wedge and the tax rate will *not* be constant anymore across states and dates.

5 Fiscal policies over states and dates

The goal of the rest of the paper is to understand the short- and long-run dynamics of optimal taxes, government consumption, and debt under doubts about the model. Proposition 5 instructs us to focus on the dynamics of the excess burden of taxation Φ_t.

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}. In particular, the excess burden of taxation increases ($\Phi_{t+1} > \Phi_t$) if 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 ($\Phi_{t+1} < \Phi_t$), if 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 happen because debt in marginal utility units has an additional *price* effect that the policymaker is manipulating in order to make debt less costly and increase the revenue from new debt issuance. With doubts about the model, an increase of the state-contingent position $z'_{s'}$ at s' decreases utility and, as a result, it increases the probability that the household assigns to this state of the world, according to (18). This is an outcome of the endogenous pessimistic expectations of the household and leads to a higher price of the respective Arrow security. The 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 return on debt falls) and harmful in the opposite case. The law of motion of Φ then just says that the planner should increase distortions (in the sense of the excess burden of taxation) for states of the world

[12] The initial period problem that is stated in the Appendix delivers the respective initial period wedges for the corresponding Φ_0.

next period for which it is cheaper to issue debt and decrease distortions for states of the world next period where debt is relatively expensive.[13]

An equivalent, more intuitive interpretation is available if we thought in terms of the policy instrument of the planner, instead of allocations. High tax rates decrease the utility of the household and increase therefore equilibrium prices through the pessimistic beliefs. By increasing tax rates on states of the world where z_{t+1} is high and reducing taxes on 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 government budget constraint and increases welfare.

The reason why the relevant object of interest is debt in *marginal utility* units comes from the logic of the intertemporal budget constraint of the government. The present discounted value of surpluses entails both an adjustment for model uncertainty, through the pessimistic expectations, and an adjustment for risk, through marginal utility.

Finally, note from the law of motion (27) that the excess burden of taxation is 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 is relevant only if debt in marginal utility units does vary across shocks or if it is actually necessary to issue debt. Otherwise, the planner follows the full confidence fiscal plan, as the following proposition shows.

Proposition 6. *Assume that either 1) $\Omega^\star(s,\Phi) = \Omega^\star(s',\Phi), \forall\Phi, \forall s \neq s'$ or that 2) initial debt is zero and there exists a $\bar{\Phi}$ such that $\Omega^\star(s,\bar{\Phi}) = \Omega^\star(s',\bar{\Phi}) = 0, \forall s \neq s'$. Then $\Phi_t = \bar{\Phi}$, where $\bar{\Phi}$ is the excess burden of taxation of the economy with full confidence in the model. Therefore, the allocation (c, h, g) and the respective wedges are the same as in the economy without doubts. Only equilibrium asset prices are different.*

Proof. See the Appendix. □

Corollary. *If 1) the utility function is as in (15) with $\rho = 1$ and we have <u>any</u> subutility of leisure $v(l)$ or if 2) initial debt is zero and the utility function is as in (15) with constant Frisch elasticity, then doubts about the model leave the second-best allocation and wedges unaltered.*

Proof. See the Appendix. □

In other words, if the planner in the full confidence economy runs a constant surplus in marginal utility units (which is zero in the case of a balanced budget), he still does so in an economy with doubts about the probability model of technology shocks and chooses the same policies (τ, Λ).

To summarize, we found sufficient conditions that *mute* the effect of model uncertainty on the Ramsey plan. For our parametric example, this happens for the knife-edge cases of $\rho = 1$ or zero

[13] 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.

initial debt and a constant Frisch elasticity.[14] In our numerical analysis, we have positive initial debt, $\rho \neq 1$ and constant Frisch elasticity. These are necessary conditions for doubts about the model to kick in.

5.2 Cyclicality of distortions and IES

We have shown that distortions covary positively with debt in marginal utility units but we still have not associated z with the cycle. To derive analytical results, we consider a two-period version of our model, where z simplifies to surplus in marginal utility units, Ω.

Without loss of generality, assume that the shocks take two values, $s_L < s_H$. Let subscripts $i = L, H$ denote the state of the world at $t = 1$. The excess burden of taxation at $t = 1$ is

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

With doubts about the model the behavior of Ω across shocks is not clear, even for our parametric example, since –in contrast to proposition 3– both the tax rate and the government share vary across shocks. Nevertheless, we can determine how $\Phi_i, i = L, H$ varies by using as a guide the Ω that pertains to the *full* confidence analysis ($\sigma = 0$), $\Omega_i^{\sigma=0}, i = L, H$. The following proposition shows that for *small* doubts about the problem, $\Omega_i^{\sigma=0}$ determines the cyclicality of distortions.

Proposition 7. *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 (15) with constant Frisch elasticity and assume that we have surpluses at $t = 1$ for $\sigma = 0$. Then, for small σ, distortions are countercyclical ($\Phi_H < \Phi_L$) if $\rho > 1$ and procyclical ($\Phi_H > \Phi_L$) if $\rho < 1$.

Proof. See the Appendix. □

Therefore, the intertemporal elasticity of substitution $1/\rho$, controls the allocation of distortions over the cycle when there are doubts about the probability model of technology shocks. Tax rates follow the same pattern as the excess burden of taxation, so they are countercyclical for an IES smaller than unity and procyclical for an IES larger than unity. Hence, when the IES is smaller than unity, "austerity" measures become *optimal*: taxes increase in bad times and decrease in good times, amplifying therefore 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.

[14] Note that in the case of *exogenous* government expenditures it would be rare to obtain a surplus in marginal utility units that does not vary across shocks, even in the knife-edge case of logarithmic utility. The reason is that the share of government consumption in output is typically not constant across shocks.

In a sense, Ω, and in infinite horizon z, captures *part* of the present value of surpluses (in terms of consumption at the first period), the other part coming from the pessimistic beliefs of the household. Surpluses and debt may be procyclical, but the countercyclicality of marginal utility –solely controlled by the IES in our homothetic specification– may be large enough to offset their cyclicality, leading to countercyclical distortions in the ambiguity setup. This role of the IES in determining the cyclicality of z was also present in environments without ambiguity. It did not lead though to any kind of "austerity" policies, due to the absence of any notion of endogenous pessimistic beliefs.

The response of the share of government consumption Λ to changes in Φ is more nuanced, since it depends on the substitutability of g with private consumption. We will say more about it in the numerical section, where we will not confine ourselves to small doubts about the model or a two-period economy.

5.3 Drifts over time and IES

Ambiguity aversion imparts also *drifts* to the excess burden of taxation, which are non-existent in full confidence economies. As in more general environments with recursive utility like Karantounias (2013b), the inverse of Φ_t is a martingale with respect to the worst-case measure 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 over time with respect to the reference model π though depends on the conditional covariance of the household's worst-case beliefs with Φ_t, since, by using the submartingale result, we get

$$E_t\Phi_{t+1} \geq \Phi_t - Cov_t(m_{t+1}, \Phi_{t+1}). \tag{28}$$

The pessimistic household assigns low probability to good times (high technology shocks) and high probability to bad times (low technology shocks). Thus, we expect m_{t+1} to be countercyclical. Given the analysis in the previous section, Φ_{t+1} is typically countercyclical if the IES is smaller than unity and procyclical if the IES is larger than unity. Therefore, we expect two different cases regarding the long-term properties of optimal policy according to π:

- *Front-loading* of distortions when $IES < 1$ ($\rho > 1$): with countercyclical distortions (the "austerity case"), we have $E_t m_{t+1}\Phi_{t+1} > E_t\Phi_{t+1}$, since bad times, which are weighed more according to the pessimistic household, bear high distortions. Consequently, the covariance term in (28) is positive, a fact which opens the possibility of a negative drift in distortions, $E_t\Phi_{t+1} \leq \Phi_t$. If this is actually true, the planner *both* increases taxes in bad times, *and* decreases on average taxes over time (according to 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 $IES > 1$ ($\rho < 1$): with procyclical distortions we have $E_t\Phi_{t+1} \geq E_t m_{t+1}\Phi_{t+1}$, since now the low-taxes states are weighed heavily from the cautious household. The covariance in (28) is negative and therefore there is a positive drift in Φ_t with respect to the reference model. This back-loading of distortions implies that the tax rate and government debt increase on average over time.

These two cases will be confirmed in the numerical simulations of the next section.

6 Numerical simulations and balanced budgets

The balanced budget policy is of particular interest in the analysis of the dynamics of optimal policy. Our economy 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 (27). If this point is ever reached, i.e. if government debt becomes zero, then it is optimal to run a balanced budget forever. In other words, Φ^* is an absorbing state. At Φ^*, doubts about the probability model affect only asset prices, as shown in proposition 6.

We proceed to the numerical treatment of the problem by treating the excess burden of taxation as a state variable with law of motion (27) and approximating the equilibrium around Φ^*.[15] All details of the approximation are relegated to the Appendix.[16] 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 of taxation takes the form

$$\Phi_{j|i}(\Phi) \simeq \Phi^* + \Phi'_{j|i}(\Phi^*)(\Phi - \Phi^*), \quad i,j = 1, ..., N. \tag{29}$$

The algorithm in the appendix delivers $\Phi'_{j|i}(\Phi^*)$, which stands for the slope of the law of motion, evaluated at the balanced budget, when we have the transition from i to j. With full confidence in the model, $\Phi'_{j|i}$ are identically equal to unity for all i, j. With doubts about the model, (29) implies that the excess burden of taxation increases over time, $\Phi_{j|i}(\Phi) > \Phi$, if $\Phi'_{j|i}(\Phi^*) > 1$. In

[15] The global solution of the problem is non-trivial due to the presence of the value functions in the constraints. See Karantounias (2013b).

[16] This approximation is not the same as an expansion around the deterministic steady state and is of interest on its own. It is similar in spirit to the approximation that Bhandari et al. (2013) employ. The approximation respects negative or positive drifts in Φ_t with respect to π and exhibits the martingale property with respect to the worst-case measure (we prove this property in the Appendix). Instead, a standard approximation around the deterministic steady state would deliver random-walk dynamics with respect to π.

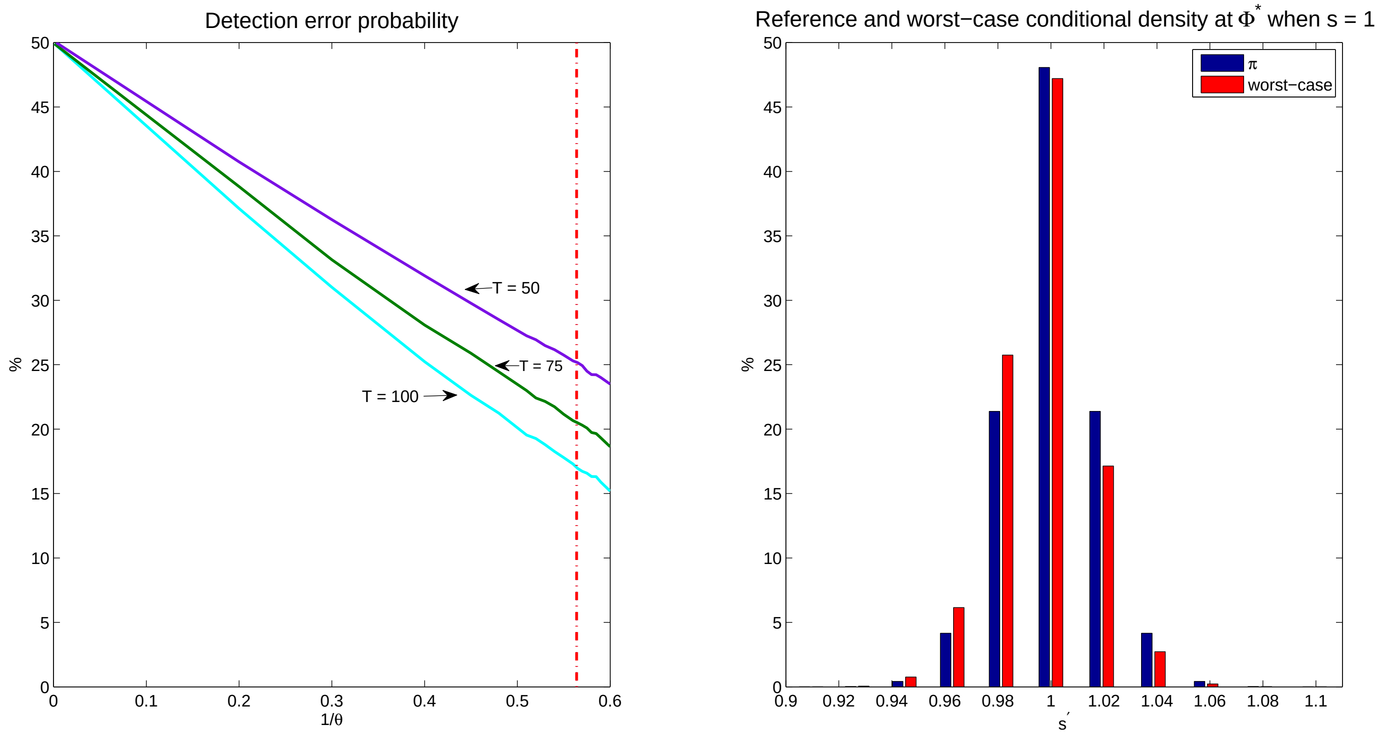

Figure 2: The left graph depicts detection error probabilities as function of $1/\theta$ for $(\rho, \psi) = (2, 1)$. The longer the length of the sample, the smaller the detection error probability for a given level of σ. For each θ $100,000$ sample paths were generated according to the reference and the worst-case model. The vertical dotted line at $\sigma = -0.564$ corresponds to $p = 25\%$ for $T = 50$, $p = 20.64\%$ for $T = 75$ and $p = 17.06\%$ for $T = 100$. The right graph depicts the respective worst-case and reference conditional density at the balanced budget, when the current shock is unity.

the opposite case of $\Phi'_{j|i}(\Phi^*) < 1$, we have $\Phi_{j|i}(\Phi) < \Phi$. Furthermore, the relative size of the coefficients matters: if $\Phi'_{\kappa|i}(\Phi^*) > \Phi'_{l|i}(\Phi^*)$ for some κ, l, then $\Phi_{\kappa|i}(\Phi) > \Phi_{l|i}(\Phi)$. These results hold when we have an excess burden of taxation that is higher than its balanced-budget value, $\Phi > \Phi^*$, which is typically the case when we start with positive initial debt.

6.1 Calibration

Our choices of parameter values are largely similar to the ones we made in the full-confidence economy. We use the same (β, ϕ_h) and the same AR(1) specification for the technology shocks. 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.

In contrast to the analysis without model uncertainty, 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 our baseline analysis we abstract from any variation at the government consumption margin and set $\psi = 1$, which implies a zero public wedge. We explore later the implications of model uncertainty on Λ. As previously, we set (α, a_h) so that the government consumption share is 20% at the first best and the household works 40% at the first-best when shocks take the value unity. The initial conditions are the same, $(s_0, b_0) = (1, 0.2)$.

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).[17] Since σ is effectively a penalty parameter, it is difficult to interpret its size. Furthermore, the worst-case beliefs of the household depend on continuation utilities as shown by (18), and therefore the effects of σ on the worst-case assessments of the household are not independent of the rest of the parameters, as long as they affect the size of V_{t+1} across shocks. To overcome these issues and to constrain interest in models that are difficult to distinguish with a limited amount of data, Hansen and Sargent propose to treat σ as a context-specific parameter, that needs to be calibrated in conjunction with the associated detection error probability.

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, the easier it is to statistically distinguish them, and the lower the detection error probability. In particular, 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

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

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 2 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 $\sigma = -0.564$ that corresponds to a detection error probability of 25% when $T = 50$, or 20.64% and 17.06% for sample paths of 75 or 100 years length respectively. Our choice of σ does not imply large doubts about the model; Hansen and Sargent (2008) regard a detection error probability as low as 10% as justifiable.

In order to get a sense about the respective pessimistic beliefs, we also plot the worst-case conditional density at the balanced budget at the right panel of figure 2, where we can clearly see how the pessimistic household shifts probability mass to low technology shocks. For the case

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

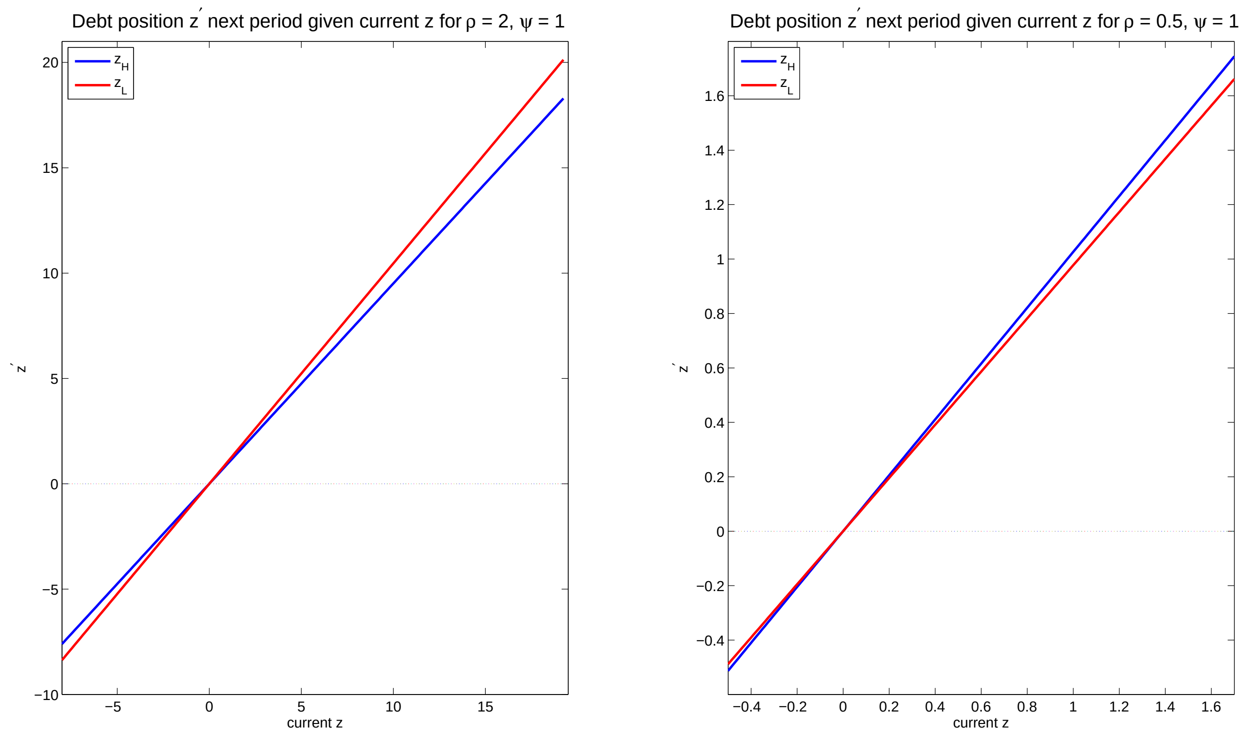

Figure 3: The left graph depicts the policy functions for z' for a low and high shock in the case of an IES that is smaller than unity. The right graph considers the case of an IES that is larger than unity. Since the government issues debt for $t \geq 1$, the relevant parts of the policy functions are for $z > 0$.

of high IES, $\rho = 0.5$, we re-calibrate σ to -2.85. This value corresponds to a detection error probability of 35% when $T = 50$.[18] If we actually kept σ to the value of -0.564 as for the low IES case, the detection error probability would be 46.94% for $T = 50$, implying that the two models are extremely difficult to distinguish. This is an obvious example of the difficulties in the interpretation of σ without the relevant context.

6.2 Policy functions and correlations

As in the two-period analysis of section 5.2, we find that distortions (in the sense of the excess burden of taxation) are countercyclical for the low IES case ("the austerity" case) and procyclical for the high IES case. In particular, the coefficients $\Phi'_{j|i}(\Phi^*)$ of the law of motion (29) are decreasing in j for all current shocks i when $\rho > 1$. In contrast, they are increasing in j for all i when $\rho < 1$. Taxes are a monotonic function of the excess burden of taxation, and, therefore, exhibit the same behavior.

To see the same result in terms of z, figure 3 plots the policy functions for debt in marginal

[18] 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, due to the possibility of multiple solutions. This is why we refrained from trying to reach a detection error probability of 25% as in the low IES case.

Table 2: Correlations of the excess burden of taxation and the tax rate.

	$\rho = 2$	$\rho = 0.5$
Correlation of $\Delta\Phi$ with s	-0.5615	0.5998
Correlation of $\Delta\Phi$ with y	-0.5538	0.6010
Autocorrelation of Φ	0.9963	0.9954
Correlation of $\Delta\tau$ with s	-0.5624	0.6001
Correlation of $\Delta\tau$ with y	-0.5549	0.6013
Autocorrelation of τ	0.9963	0.9954

The table depicts median 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.

utility units for the case of low and high IES. As expected, the value of debt is countercyclical when the IES is low and procyclical when the IES is high. Table 2 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 4 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.

6.3 Long-run dynamics

As in the case with full confidence in the model, the Ramsey plan prescribes a deficit in the initial period. In the subsequent periods though, the planner is running either a decreasing or an increasing surplus-to-output ratio, depending on the IES. In the first case, the planner is trying to repay the entire stock of debt. In the second case, the planner is postponing distortions to the future and increases public debt over time.

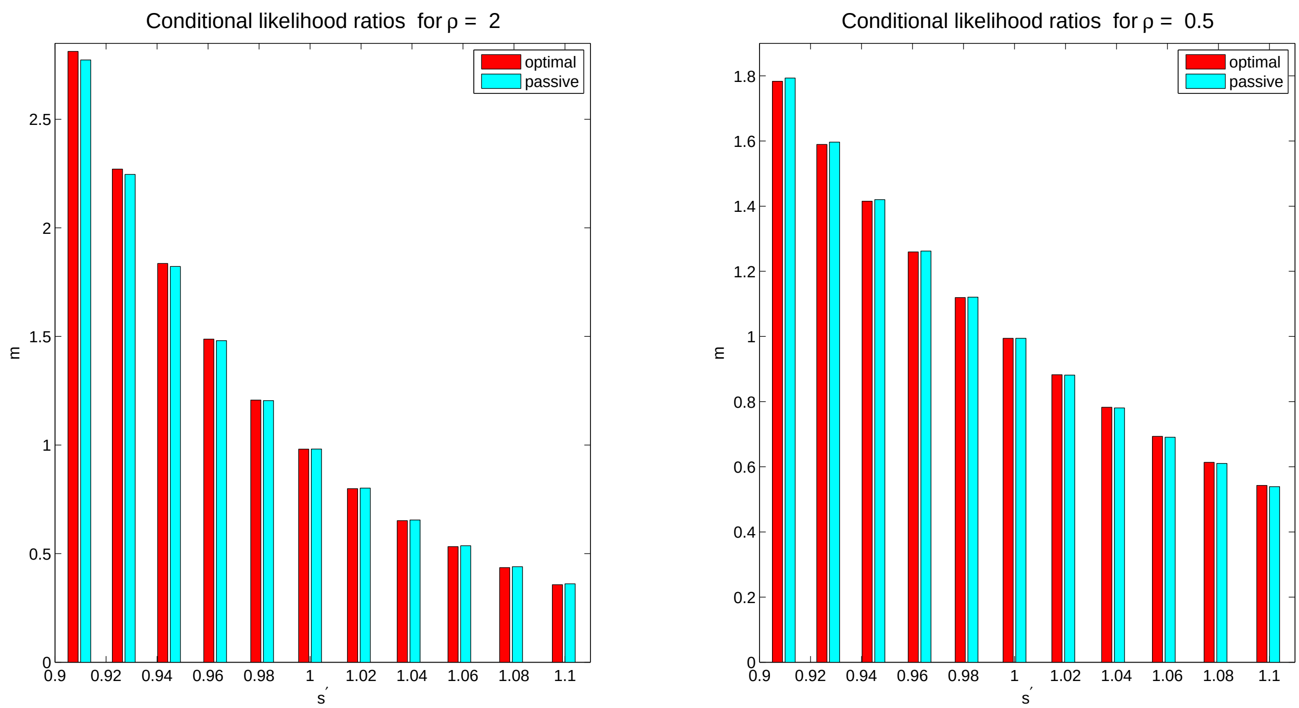

Figure 4: Amplifying pessimistic expectations for IES smaller than unity (left graph) and mitigating pessimistic expectations for IES larger than unity (right graph). The current shock is $s = 1$.

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

For the baseline calibration of $\rho = 2$, we find that $\sum_j \pi(j|i)\Phi'_{j|i}(\Phi^*) < 1, \forall i$. The approximate law of motion (29) implies then a *negative* drift with respect to π, $\sum_j \pi(j|i)\Phi_{j|i} < \Phi$, when $\Phi > \Phi^*$. Figure 5 displays a typical sample path that captures the front-loading of distortions. The optimal plan features decumulation of debt and a balanced primary budget in the long-run. 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 6 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

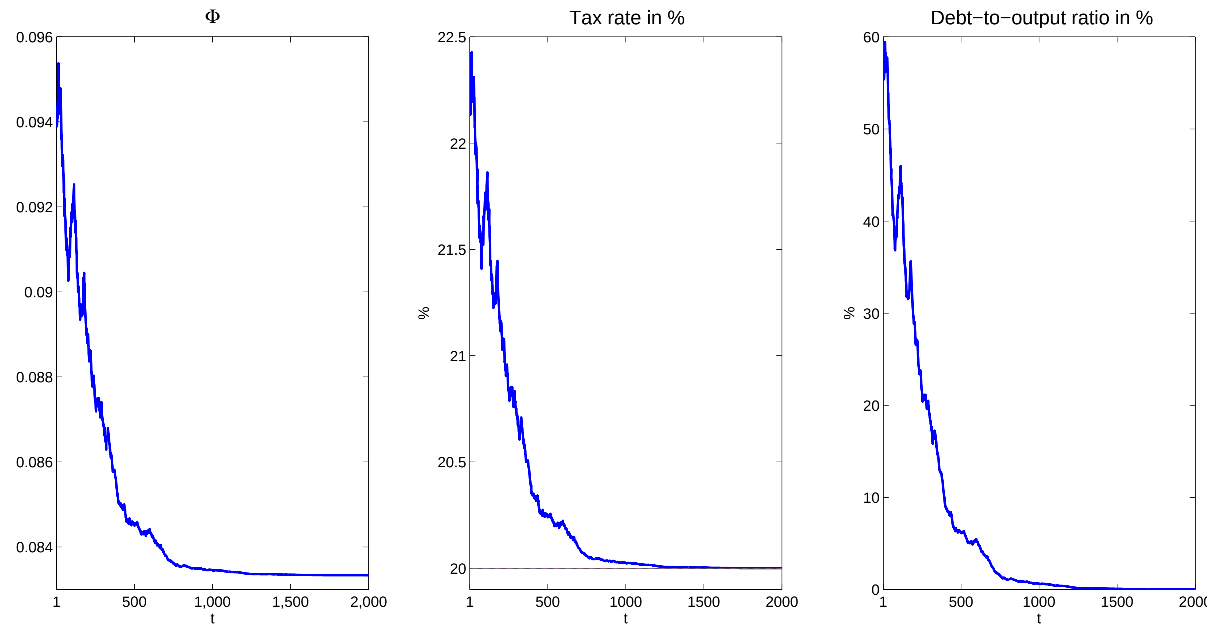

Figure 5: Typical sample path for $\rho = 2$. It displays long-run convergence to the balanced budget and zero public debt for $\rho = 2$. The balanced budget tax rate is 20%. The government runs at $t = 0$ a deficit that is 6.05% of output.

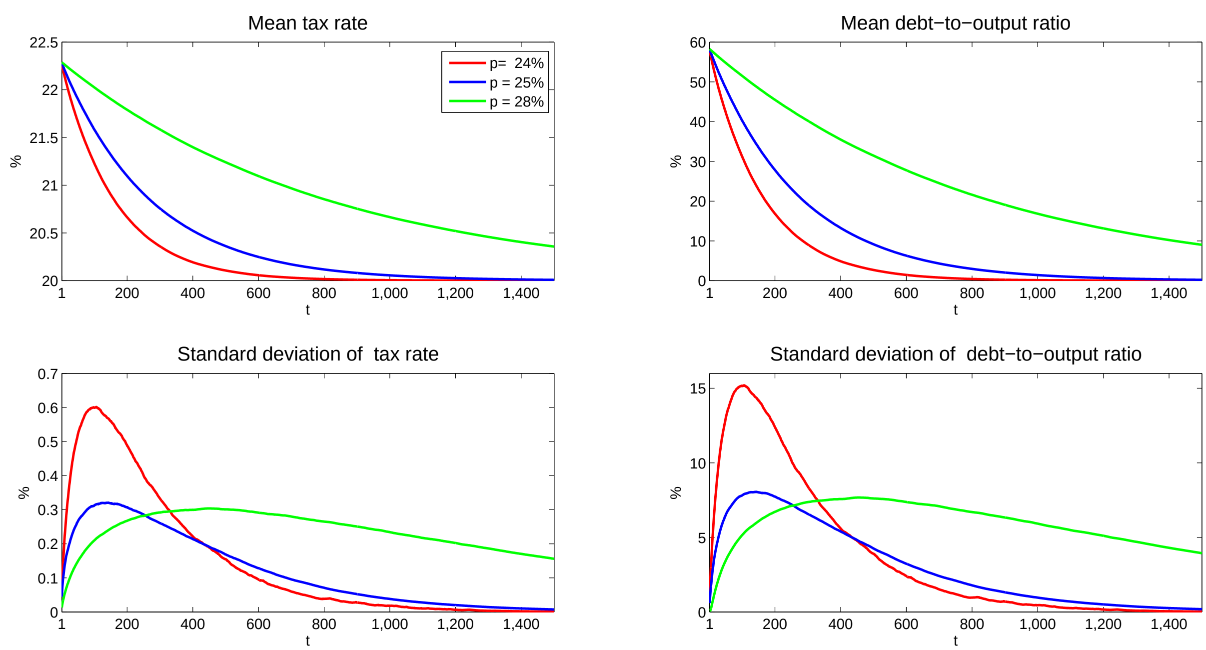

Figure 6: Moments of the tax rate and the debt-to-output ratio over time for $\sigma \in \{-0.5868, -0.564, -0.492\}$. These values of σ imply a p of 24%, 25% and 28% respectively. 10,000 sample paths were used for each σ.

non-monotonic way over time, featuring a hump-shape pattern. 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

Table 3: Median half-time in years of debt-to-output ratio.

	$\sigma_a = 0.03$	$\sigma_a = 0.04$	$\sigma_a = 0.05$
$\rho_a = 0.95^4$	173	90	50
	(25%)	(18.46%)	(12.99%)
$\rho_a = 0.98^4$	102	50	25
	(19.8%)	(12.96%)	(8.17%)
$\rho_a = 0.99^4$	88	42	20
	(18.42%)	(11.73%)	(7.07%)

$10,000$ sample paths were simulated for each specification (ρ_a, σ_a). We report the median period at which the debt ratio reaches $0.5 \cdot b_1/y_1$. The average b_1/y_1 is $56.3\% - 58.11\%$ depending on the specification. The conditional standard deviation is adjusted for each ρ_a so that the unconditional volatility remains the same for each column. The parameter σ was kept at its baseline value, $\sigma = -0.564$. We also report in parentheses the respective detection error probability p for $T = 50$.

it reached its stationary counterpart, $\frac{\sigma_\epsilon}{\sqrt{1-\phi^2}}$. 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. Furthermore, 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.

Debt reduction and the shock process. The speed of debt reduction depends crucially on the stochastic properties of the shocks. Table 3 reports the median halftime of the debt-to-output ratio for different persistence-volatility pairs (ρ_a, σ_a) of the shocks s. Increased persistence for a given unconditional volatility, or increased volatility for a given persistence, reduce the median halftime considerably. This result comes from the fact that highly volatile and persistent shocks are inviting the planner to manage the pessimistic expectations more aggressively, speeding up the convergence to the balanced budget. For example, for our baseline calibration the adjustment is slow, implying a median halftime of 173 years, whereas for a high (ρ_a, σ_a) pair it is reduced to just 20 years. The aggressive expectation management and the resulting fiscal policies lead to a worst-case model that differs in a non-trivial way from the reference model. As a result it becomes easier to statistically distinguish between the two models when the persistence or volatility are increased, leading to smaller detection error probabilities, as table 3 attests.

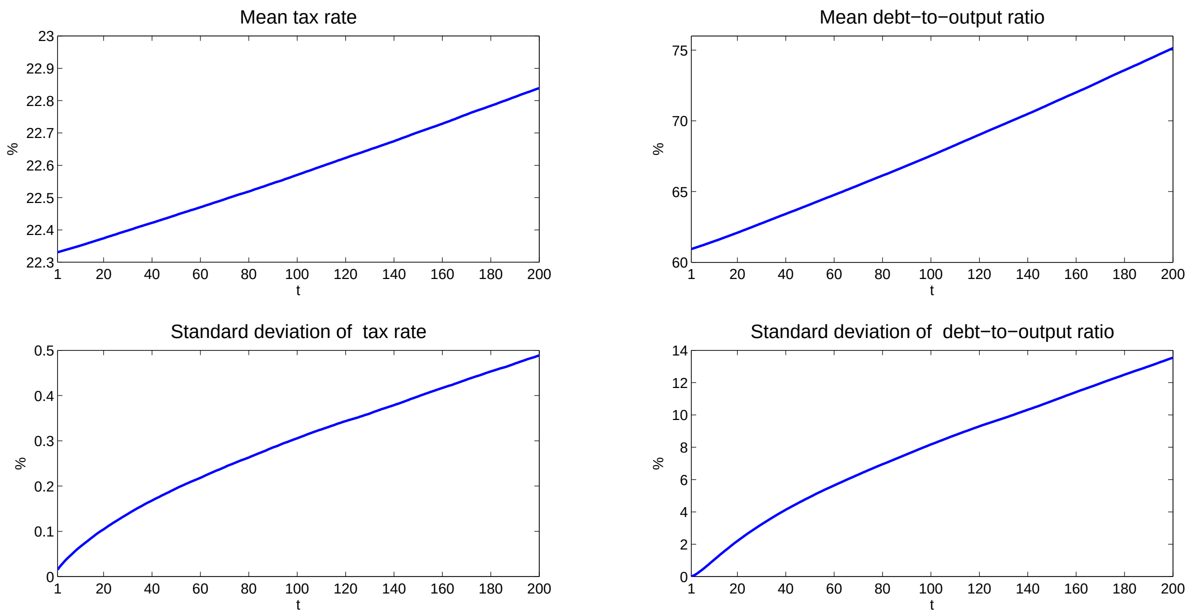

Figure 7: Mean and standard deviation of the tax rate and the debt-to-output ratio when $(\rho, \psi) = (0.5, 1)$. Doubts about the model are set to $\sigma = -2.85$, that corresponds to a detection error probability of 35% for $T = 50$.

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

We showed before that we have a procyclical allocation of distortions when the IES is larger than unity. Furthermore, we find that $\sum_j \pi(j|i)\Phi'_{j|i}(\Phi^*) > 1, \forall i$, which implies a *positive* drift with respect to π, when $\Phi > \Phi^*$. Figure 7 displays the mean and the standard deviation of the tax rate and the debt-to-output ratio. The positive drift in the tax rate is reflected in a positive drift in the debt-to-output ratio.[19] The intuition is similar as before: since good times happen more often according to the data-generating process than what the pessimistic household thinks, and since now good times bear higher taxes than bad times, average tax rates (and debt) increase over time.

6.4 Government consumption

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 4 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

[19]We refrain from reporting moments for the entire length of our samples for this case, because the approximation becomes progressively worse since the balanced budget Φ^* is not a stable point anymore. In contrast, the approximation becomes progressively better when $IES < 1$.

Table 4: Correlation of $\Delta\Lambda$ with the technology shock.

	Substitutes ($\psi = 0.5$)	Complements ($\psi = 2$)
Low IES ($\rho = 2$)	0.4468	-0.5802
High IES ($\rho = 0.5$)	-0.6002	0.4507

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 median statistics across sample paths.

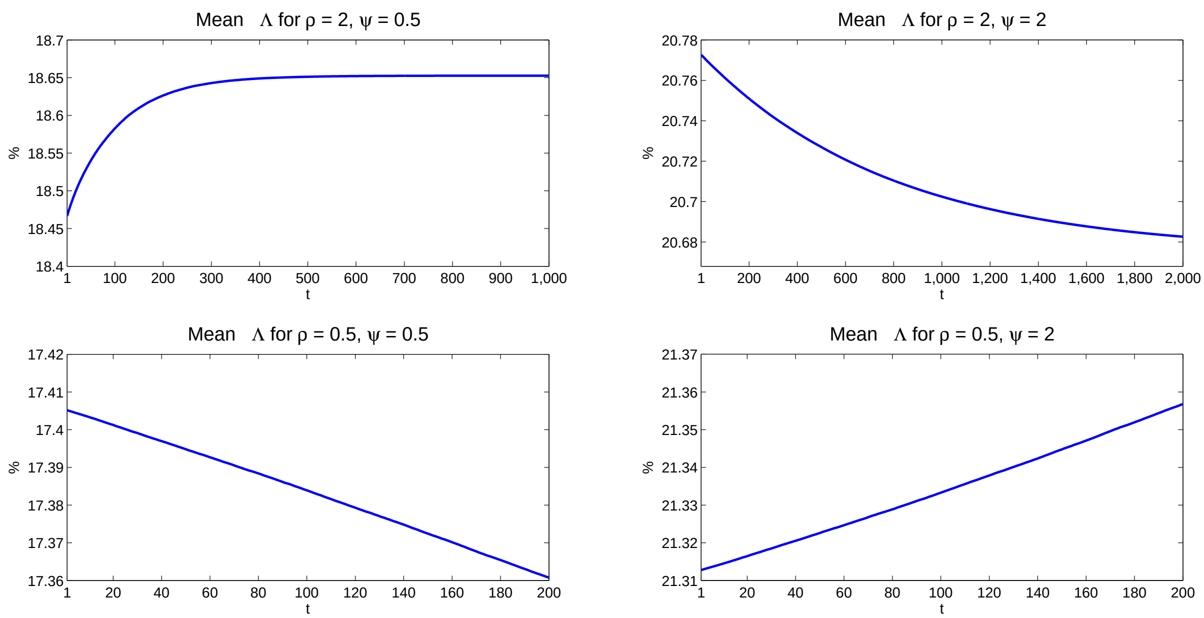

Figure 8: Evolution of the mean government share over time. The two graphs on the top consider the case of $\rho = 2$. The two graphs on the bottom consider the case of $\rho = 0.5$. Graphs on the left correspond to the substitutes case ($\psi = 0.5$) and graphs on the right to the complements case, ($\psi = 2$). 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%.

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 4 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 8 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 $\psi = 1$.[20]

7 Concluding remarks

In this paper we found that two, diametrically opposite, policies are optimal when there is ambiguity about the cycle: "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.[21]

In our study we abstracted from any kind of under-utilization of resources or any kind of default risk, that may annul or favor fiscal consolidation arguments.[22] 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. When this pessimism is unfounded, and good, low-tax times happen relatively often, governments manage to reduce debt over time and reach balanced budgets in the long-run, where pessimism ceases to be relevant for optimal policy.

[20] 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. However, this would shift the focus away from ambiguity about the cycle and strain the interpretation – does the agent doubt his own model of preference shocks?

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

[22] 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

$$
\begin{aligned}
c_t(s^t): \quad & U_c(s^t) + \Phi\Omega_c(s^t) = \lambda_t(s^t) && \text{(A.1)} \\
h_t(s^t): \quad & -U_l(s^t) + \Phi\Omega_h(s^t) = -\lambda_t(s^t)s_t && \text{(A.2)} \\
g_t(s^t): \quad & U_g(s^t) + \Phi\Omega_g(s^t) = \lambda_t(s^t) && \text{(A.3)}
\end{aligned}
$$

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

$$
\begin{aligned}
c_0: \quad & U_{c0} + \Phi(\Omega_{c0} - U_{cc0}b_0) = \lambda_0 && \text{(A.4)} \\
h_0: \quad & -U_{l0} + \Phi(\Omega_{h0} + U_{cl0}b_0) = -\lambda_0 s_0 && \text{(A.5)} \\
g_0: \quad & U_{g0} + \Phi(\Omega_{g0} - U_{cg0}b_0) = \lambda_0 && \text{(A.6)}
\end{aligned}
$$

Eliminate now the multiplier λ_t from (A.1), (A.2) and (A.3) to get (13) and (12). Use (A.4)-(A.6) to get the respective wedges for $t = 0$:

$$
\begin{aligned}
\frac{U_{g0} + \Phi(\Omega_{g0} - U_{cg0}b_0)}{U_{c0} + \Phi(\Omega_{c0} - U_{cc0}b_0)} &= 1 && \text{(A.7)} \\
\frac{U_{l0} - \Phi(\Omega_{h0} + U_{cl0}b_0)}{U_{c0} + \Phi(\Omega_{c0} - U_{cc0}b_0)} &= s_0 && \text{(A.8)}
\end{aligned}
$$

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.2 Proof of proposition 2

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

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

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

To prove proposition 2 express (12) 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 (14) 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 (13) 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.8) and (A.7) 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.3 Proof of proposition 3

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

$$A\kappa^{-\psi} = 1 + \frac{\Phi(1-\psi)}{1 + \Phi(\psi - \rho)[1 + A\kappa^{1-\psi}]^{-1}}. \tag{A.9}$$

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.9) 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.9) 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.9) 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 $\tau'(0) = \epsilon_{cc}(\kappa^{FB}) + \phi_h > 0$.

2c) 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 output $y = sh$. The function $H(.)$ is

$$H(\Phi) \equiv \left[\frac{1-\tau}{a_h}\frac{(1-\alpha)\left(\frac{c}{u}\right)^{\rho-\psi}}{(1-\Lambda)^{\rho}}\right]^{\frac{1}{\rho+\phi_h}}$$

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 $\kappa(\Phi)$.

2d-e) 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 (as expressed in **2c**) 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) using the same argument as in **2c**. The result for the derivative of $\Omega^\star$ follows.

A.4 Proof of proposition 4

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 (26) with the same qualification.

B.2 Proof of proposition 6

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 (27) 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 $U_{c0}b_0 = \Omega_0 + \beta\Omega^*/(1-\beta)$. 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. 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 (27). 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 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 $\bar{\Phi}$.

B.3 Proof of corollary to proposition 6

1) We have already shown in (16) that the surplus in marginal utility units does not vary across shocks for $\rho = 1$ when we have a constant Frisch elasticity, so the first assumption of proposition 6 applies. We will now 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 (13) 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 4. 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 7

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 $\Phi_i'(0) = \Phi_0'(0) - \Phi(0)^2\eta_i(0)$.[23] 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 $\Omega_H^{\sigma=0} < \Omega_L^{\sigma=0}$.

[23] 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, $\Phi = 0$.

2) With full confidence in the model Φ is constant, and therefore the surplus in marginal utility units is countercyclical if $\rho > 1$ and procyclical if $\rho < 1$ according to **e)** of proposition 3. The result follows.

C Balanced budget approximation

C.1 Laws of motion

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 $s_t = s_i$. 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

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

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.

Laws of motion. Let the current shock be i and the current excess 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 (27) implies the system $\mathbf{F_i} = \mathbf{0}$, where $\mathbf{0}$ 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 (29). 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

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

Before we proceed to the calculation of the matrix $J^*_{\eta_i}$, several remarks are in order. At first, (C.1) shows that for $\sigma = 0$ we have $\Phi'_{j|i}(\Phi^*) = 1 \forall i, j$, so the approximate law of motion (29) implies that $\Phi_{j|i}(\Phi) = \Phi \forall i, j$, as expected. Second, we have the following lemma:

Lemma C.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

$$\sum_j \pi(j|i) m^*_{j|i} \Phi'_{j|i}(\Phi^*) = 1, \forall i \tag{C.2}$$

If (C.2) holds, then the approximate law of motion (29) implies that $\sum_j \pi(j|i) m^*_{j|i} \Phi_{j|i}(\Phi) = \Phi$ and the result follows. To show (C.2) 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

$$\sum_j \pi(j|i) m^*_{j|i} \big[\sum_k \frac{\partial \eta^*_{j|i}}{\partial \Phi_k} \Phi'_{k|i}(\Phi^*)\big] = 0, \quad \text{or} \quad e_i^T \Pi^* J^*_{\eta_i} \cdot \begin{pmatrix} \Phi'_{1|i}(\Phi^*) \\ \vdots \\ \Phi'_{N|i}(\Phi^*) \end{pmatrix} = 0, \forall i, \tag{C.3}$$

where e_i the vector with unity at position i and zero otherwise. Pre-multiply system (C.1) 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 (C.3) *zero*, a fact which delivers ultimately (C.2).

$\square$

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

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

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 (C.1), we need the sensitivity of the debt positions with respect to the excess burden of taxation $\mathbf{z}'$.

Proposition 6 showed the necessity of variation in debt positions across states of the world for a non-constant allocation of distortions. The respective result in a first-order approximation around the balanced budget concerns the sensitivity of debt positions $z'_i(\Phi^*)$. In particular, if $z'_i(\Phi^*)$ are the *same* across shocks, then $\Phi'_{j|i}(\Phi^*) = 1, \forall i, j$. To see that clearly, let $z'_i(\Phi^*) = \bar{Z}, \forall i$. System C.1 simplifies to

$$\Big[I+\sigma\bar{Z}\cdot(\Phi^*)^2\big(I-\mathbf{1}\cdot(e_i^T\Pi^*)\big)\Big]\begin{pmatrix}\Phi'_{1|i}(\Phi^*)\\ \vdots\\ \Phi'_{N|i}(\Phi^*)\end{pmatrix}=\mathbf{1}$$

which, by collecting terms and rearranging, can be rewritten as

$$(1+\sigma\bar{Z}\cdot(\Phi^*)^2)\begin{pmatrix}\Phi'_{1|i}(\Phi^*)\\ \vdots\\ \Phi'_{N|i}(\Phi^*)\end{pmatrix}=\mathbf{1}\Bigg(1+\sigma\bar{Z}\cdot(\Phi^*)^2e_i^T\Pi^*\cdot\begin{pmatrix}\Phi'_{1|i}(\Phi^*)\\ \vdots\\ \Phi'_{N|i}(\Phi^*)\end{pmatrix}\Bigg)=\big(1+\sigma\bar{Z}\cdot(\Phi^*)^2\big)\mathbf{1}\text{ by (C.2).}$$

So we finally get $\Phi'_{j|i}(\Phi^*)=1\forall i,j$, for σ small enough so that $1+\sigma\bar{Z}\cdot(\Phi^*)^2\neq 0$.

C.2 Implementability constraints

In order to calculate $z_i'(\Phi^*)$ we need the implementability constraint

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

Differentiate implicitly with respect to Φ to get

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

which at Φ^* becomes

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

Define the $N\times N$ matrix of the derivatives of the excess burden of taxation

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

We can write the differentiated implementability constraints as

$$\mathbf{z}' = \mathbf{\Omega}' + \beta(\Pi^* \circ \mathbf{\Phi})\mathbf{z}' \Rightarrow \mathbf{z}' = \Big(I - \beta(\Pi^* \circ \mathbf{\Phi})\Big)^{-1}\mathbf{\Omega}', \tag{C.5}$$

where $\mathbf{\Omega}' \equiv [\Omega_1'(\Phi^*), ..., \Omega_N'(\Phi^*)]^T$. Note that as long as $\Phi'_{j|i}(\Phi^*)$ are positive, (C.2) implies that $\Pi^* \circ \mathbf{\Phi}$ is a stochastic matrix and therefore (C.5) is a modified presented discounted value formula. The derivatives of the surplus in marginal utility units at Φ^* are $\Omega_i'(\Phi^*) = \big(\tau'(\Phi^*) - \Lambda'(\Phi^*)\big)U_c(i, \Phi^*)y(i, \Phi^*)$. As shown earlier, output in marginal utility units is procyclical if $\rho < 1$ and countercyclical if $\rho > 1$. In all of our calibrations $\tau'(\Phi^*) > \Lambda'(\Phi^*)$.

Solution. We have $N^2 + N$ unknowns ($\Phi'_{j|i}(\Phi^*)$ and $z_i'(\Phi^*)$) and $N^2 + N$ equations from (C.1) and (C.5). To solve for the unknowns we proceed as follows:

- Make a guess for $\mathbf{\Phi}$. Derive induced derivatives of the relative debt positions $\mathbf{z}'$ from (C.5).
- Use $\mathbf{z}'$ to get the Jacobian $J^*_{\eta_i}, \forall i$ from (C.4) and update the guess for $\mathbf{\Phi}$ by solving the systems (C.1).
- 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 the linear approximation around Φ^* only for the excess burden of taxation and for the debt in marginal utility units z. Given the value of Φ from the approximate law of motion, we use the non-linear functions for (τ, Λ) and (c, h, g), so our method is "hybrid".

Detection error probability. In order to calculate detection error probabilities we need to simulate according to the worst-case model. It is easy to see that the derivatives $V_i'(\Phi^*), i = 1, .., N$ are a solution to the linear system

$$\begin{pmatrix} V_1'(\Phi^*) \\ \vdots \\ V_N'(\Phi^*) \end{pmatrix} = \Big(I - \beta(\Pi^* \circ \mathbf{\Phi})\Big)^{-1} \begin{pmatrix} U_1'(\Phi^*) \\ \vdots \\ U_N'(\Phi^*) \end{pmatrix}.$$

The utility of the household that determines the worst-case model is given by the approximation $V_i(\Phi) \simeq V_i(\Phi^*) + V_i'(\Phi^*)(\Phi - \Phi^*), \forall i$.

References

Ambler, Steve and Alain Paquet. 1996. Fiscal spending shocks, endgoenous government spending, and real business cycles. *Journal of Ecnomic 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. Mimeo, Federal Reserve Bank of Minneapolis.

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. 2013. Taxes, debts and redistributions with aggregate shocks. Mimeo, New York University.

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).

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

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.

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.

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. Mimeo, Federal Reserve Bank of Atlanta.

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

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.

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. 2015. Sovereign Default Risk and Uncertainty Premia. Mimeo, University of California at Berkeley.

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.

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