Fiscal austerity in ambiguous times *

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

January 17, 2018

Online Appendix

(not for publication)

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

JEL classification: D80; E62; H21; H63.

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

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

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

A Balanced budget approximation

A.1 Preliminaries

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

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

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

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

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

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

A.2 Approximate law of motion

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

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

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

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

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

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

where

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

We are going to work under the following assumption.

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

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

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

A.3 Three lemmata

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

Proof. We will show that

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

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

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

Differentiate implicitly with respect to Φ to get

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

At Φ^* this expression simplifies to

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

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

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

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

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

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

Therefore:

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

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

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

Rewrite the above as

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

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

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

The inverse of the diagonal matrix is

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

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

□

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

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

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

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

Proof. Consider the implementability constraints

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

Differentiate implicitly with respect to Φ to get

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

which at Φ^* becomes

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

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

□

A.4 Proof of proposition 7

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

A.5 Proof of proposition 8

Write the surplus in marginal utility units as

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

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

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

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

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

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

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

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

We have then

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

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

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

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

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

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

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

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

and note that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A.6 An algorithm

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

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

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

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

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

B Government consumption share

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

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

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

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

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

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

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

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

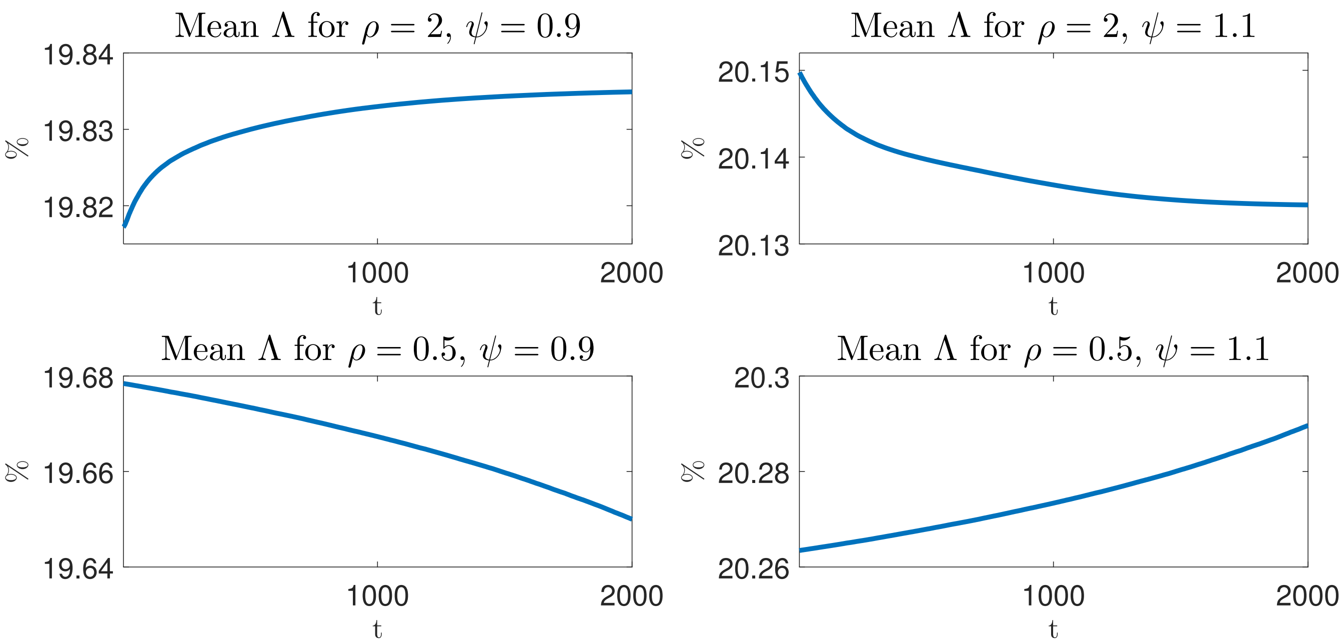

Figure 1: Evolution of the mean government share over time. The two graphs on the top consider the case of $\rho = 2$. The two graphs on the bottom consider the case of $\rho = 0.5$. Graphs on the left correspond to the substitutes case ($\psi = 0.9$) and graphs on the right to the complements case,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