ADEMU

A Dynamic Economic and Monetary Union

ADEMU WORKING PAPER SERIES

The effects of government spending endogeneity on estimated multipliers in the U.S.

Alban Moura†

November 2015

WP 2016/015

www.ademu-project.eu/publications/working-papers

Abstract

This paper uses an estimated sticky-price model to identify endogenous movements in government consumption in the U.S. economy. Two feedback effects are considered, one originating from the stock of public debt and one from contemporaneous output. The data provide significant statistical evidence in favor of such mechanisms, even though a subsample analysis reveals that their strength may have decreased over time. Monte Carlo simulations assessing a DSGE model with exogenous spending and various identified VARs suggest that failing to account for these feedbacks may induce a severe upward bias in estimated multipliers.

***Keywords**:* Government spending multiplier, endogenous fiscal policy, structural econometrics

Jel codes: C32, E62, H30

---

† Toulouse School of Economics, Universit_e Toulouse I Capitole, Manufacture des Tabacs, Bâtiment F, 21 allée de Brienne, 31000 Toulouse, France. E-mail: alban.moura@tse-fr.eu

Horizon 2020
The Framework Programme for Research and Innovation

ADEMU
A Dynamic Economic and Monetary Union

Acknowledgments

For helpful comments, I thank R. Chahrour, M. Dupaigne, C. Hellwig, M. Normandin, and especially P. Féve and F. Portier, as well as two anonymous referees and participants at the XIXth Vigo Workshop on Dynamic Macroeconomics and 2014 EEA- ESEM conference. I am also grateful to the ANR \Analyse des Multiplicateurs Fiscaux" (ANR-13-BSH1-0002) for financial support. Any remaining errors are my own. This project is related to the research agenda of the ADEMU project "A Dynamic Economic and Monetary Union, funded by the EU Horizon 2020 Program under grant agreement N° 649396 (ADEMU).

The ADEMU Working Paper Series is being supported by the European Commission Horizon 2020 European Union funding for Research & Innovation, grant agreement No 649396.

1. Introduction

Fiscal stimulus packages implemented throughout the world in response to the last recession have renewed the academic interest in the general equilibrium effects of fiscal policy, especially changes in government spending. Empirically, the difficulty in measuring such effects is to identify exogenous policy shocks, since observed movements in fiscal variables may as well reflect automatic responses to economic conditions or financing constraints. Historically, standard practice has been to put this difficulty aside by assuming that government spending is predetermined with respect the state of the economy and follows an exogenous stochastic process. This assumption underlies many empirical works on spending multipliers, both using dynamic stochastic general equilibrium (DSGE) models (Cogan, Cwik, Taylor, and Wieland, 2010; Christiano, Eichenbaum, and Rebelo, 2011) or structural vector autoregressions (SVARs; Fatás and Mihov, 2001; Blanchard and Perotti, 2002; Galí, López-Salido, and Vallés, 2007).

A strand of papers calls into question the validity of this approach by emphasizing the presence of endogenous movements in government spending. For instance, estimating real-business-cycle models, McGrattan (1994) and Jones (2002) both report positive responses of public expenditures to, respectively, contemporaneous productivity shocks and contemporaneous output. Instead, Fève, Matheron, and Sahuc (2013) obtain the opposite result that government spending is countercyclical in the U.S. While contradictory, these findings hint that the exogeneity assumption is not supported by the data. Additionally, Leeper, Plante, and Traum (2010) find a significant feedback from the level of public debt to spending, a mechanism absent from papers adopting a Ricardian framework. Eventually, endogenous links are uncovered using simple regressions by Clemens and Miran (2012), who argue that balanced-budget rules induce substantial procyclicality in state government spending in the U.S.

In this context, the objectives of this paper are twofold. First, I provide new estimates of endogenous movements in government consumption in the postwar U.S. economy using a New-Keynesian DSGE model in the spirit of Smets and Wouters (2007). I augment the baseline structure by a richer description of the fiscal sector, including explicit debt financing as in Leeper, Plante, and Traum (2010), distortionary taxation as in Uhlig (2010) and Drautzburg and Uhlig (2011), and feedback policy rules allowing fiscal instruments to respond to contemporaneous output and to the lagged level of debt. The model's strong propagation mechanisms are important for valid estimation of policy rules coefficients, since a flex-price economy with weak internal channels would attribute to feedback effects much of the unexplained comovements between, say, output and government consumption. I estimate the model using quarterly U.S. data from 1960 to 2007 with Bayesian methods, considering the behavior of a general government sector aggregating federal and state government accounts. Second, I use the estimated model as laboratory to test output multipliers derived from three standard econometric approaches: two that abstract from potential endogeneity—a DSGE model with exogenous government spending and a VAR identified by exogeneity restrictions—and one that tries to control for it—a VAR identified by Mountford and Uhlig's (2009) sign restrictions.[1]

Two results stand out. First, I find significant statistical evidence in favor of government consumption endogeneity over the 1960-2007 sample, with a significant response of spending to both the debt-to-output ratio and contemporaneous output. Digging deeper, I uncover some subsample instability in the endogeneity patterns, with a fall in the individual elasticities of government consumption to debt and output over time. Second, I show that omitting to control for endogeneity induces an upward bias in estimated output multipliers, especially so when the policy shock is identified using SVARs. Monte Carlo experiments suggest that the estimation error may be sizable, with output multipliers sometimes overestimated by a factor 2, and may be present even for a DGP featuring only moderate feedback effects. Interestingly, of the three identification approaches under test, sign restrictions have the worst performance in spite of being especially designed to handle policy endogeneity.

The paper contributes to a rising literature trying to understand the determinants of fiscal policy and their implications for the identification of fiscal multipliers. While existing DSGE models with fiscal policy rules either consider flex-price economies, abstract from debt financing, or consider only the behavior of the federal U.S. government, the model estimated here features nominal frictions that have proved important to match the data, specifies rich fiscal policy rules, and considers the behavior of the general government sector to provide a comprehensive picture of the forces shaping government consumption in the U.S. Focusing on the general government sector seems especially important, both because empirical work on multipliers typically considers aggregate public spending as a whole and because of Clemens and Miran's (2012) finding that local government spending has a strong endogenous component. The paper is also the first to quantify the effects of fiscal endogeneity on the outcomes of leading identification approaches. By doing so, it complements Chahrour, Schmitt-Grohé, and Uribe (2012) who evaluate the ability of VARs to propagate the effects of fiscal shocks but abstract from identification issues.

The rest of the paper is organized as follows. Section 2 sets up the DSGE model and discusses the specification of fiscal rules. Section 3 describes the estimation procedure and the data. It also reports estimation results and evaluates the strength of endogenous feedback effects on government consumption in different subsamples. Eventually, section 4 uses the

[1] I do not test two alternative empirical strategies: Ramey's (2011) correction for anticipated shocks and Bouakez, Chihi, and Normandin's (2014) use of conditional heteroscedasticity. Both would require nontrivial changes in the specification and the estimation of the DSGE model, namely the inclusion of news shocks and of time-varying volatility. I leave exploration of these for future work.

DSGE model to test the performance of alternative identification techniques in presence of spending endogeneity. Section 5 concludes.

2. Model

I use a New-Keynesian model with a focus on fiscal policy, similar to the one laid out in Traum and Yang (2011). Compared to most empirical monetary models, its specificity lies in the interplay of public debt, distortionary taxation, and endogenous fiscal policy rules with the government budget constraint. Such features allow for a rich description of fiscal policy, enhancing the model's ability to fit the data and to identify the dynamic effects of fiscal shocks.

2.1. **Firms.** The economy produces a single final good used for consumption, investment, and government spending, and a continuum of intermediate goods indexed by $f \in [0, 1]$. The final-good sector is perfectly competitive, while there is monopolistic competition in the markets for intermediates.

2.1.1. *Final-good firms.* The final good is produced by combining intermediate goods according to

$$Y_t = \left(\int_0^1 Y_t(f)^{\frac{1}{1+\eta_t^p}} df \right)^{1+\eta_t^p},$$

where $Y_t(f)$ is the quantity of intermediate good f used in final-good production and η_t^p is a shock to the good market markup evolving according to a stationary process:

$$\ln \frac{\eta_t^p}{\eta^p} = \rho_p \ln \frac{\eta_{t-1}^p}{\eta^p} + \epsilon_t^p, \quad \text{with } \epsilon_t^p \sim \text{iid}N(0, \sigma_p^2) \text{ and } \eta^p > 0.$$

Letting $P_t(f)$ denote the price of intermediate good f and P_t the associated price index, cost minimization implies a demand structure of the form

$$Y_t(f) = \left(\frac{P_t(f)}{P_t} \right)^{-\frac{1+\eta_t^p}{\eta_t^p}} Y_t, \quad \text{with } P_t = \left(\int_0^1 P_t(f)^{-\frac{1}{\eta_t^p}} df \right)^{-\eta_t^p}.$$

Perfect competition in the final sector implies that P_t is also the price of the final good.

2.1.2. *Intermediate-good firms.* Each intermediate good f is produced by a monopolist according to

$$Y_t(f) = u_t^a [v_t K_{t-1}(f)]^\alpha L_t(f)^{1-\alpha},$$

where $v_t K_{t-1}(f)$ denotes effective capital input taking utilization v_t into account, $L_t(f)$ denotes labor input, and $\alpha \in [0, 1]$ is the capital share. There is no fixed cost of production and u_t^a is a productivity shock evolving according to a stationary process:

$$\ln u_t^a = \rho_a \ln u_{t-1}^a + \epsilon_t^a, \quad \text{with } \epsilon_t^a \sim \text{iid}N(0, \sigma_a^2).$$

The static cost-minimization problem implies that all producers have the same nominal marginal cost

$$MC_t = \frac{1}{u_t^a}\left(\frac{W_t}{1-\alpha}\right)^{1-\alpha}\left(\frac{R_t^k}{\alpha}\right)^{\alpha},$$

where W_t and R_t^k denote the nominal wage and rental rate of capital.

Intermediate firms face Calvo (1983) frictions in nominal price setting. Each period, an intermediate firm can reoptimize its price with probability $1-\xi_p$. Those that cannot do so index their prices to lagged inflation according to

$$P_t(f) = \pi_{t-1}^{\iota_p} P_{t-1}(f),$$

where $\pi_t = P_t/P_{t-1}$. A firm that is able to reoptimize at date t solves

$$\max_P \; E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \Lambda_{t+s}(1-\tau_{t+s}) \left[\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{\iota_p} P - MC_{t+S}\right]\left[\left(\frac{P_{t+s-1}}{P_{t-1}}\right)^{\iota_p} \frac{P}{P_{t+s}}\right]^{-\frac{1+\eta_{t+s}^p}{\eta_{t+s}^p}} Y_{t+s},$$

where Λ_t is the household's stochastic discount factor for nominal payoffs and τ_t the marginal tax rate on profits. Letting $P_t^\star$ denote the solution to this maximization program, the aggregate price index evolves according to

$$\pi_t = \left[(1-\xi_p)\pi_t^{-\frac{1}{\eta_t^p}}\left(\frac{P_t^\star}{P_t}\right)^{-\frac{1}{\eta_t^p}} + \xi_p \pi_{t-1}^{-\frac{\iota_p}{\eta_t^p}}\right]^{-\eta_t^p}.$$

2.2. **Labor market.** A perfectly competitive labor packer purchases differentiated labor services supplied by households and transforms them into a composite labor input L_t usable by firms:

$$L_t = \left(\int_0^1 L_t(l)^{\frac{1}{1+\eta_t^w}} dl\right)^{1+\eta_t^w},$$

where $L_t(l)$ denotes labor service of type l and η_t^w is a shock to the labor market markup evolving according to a stationary process:

$$\ln\frac{\eta_t^w}{\eta_w} = \rho_w \ln\frac{\eta_{t-1}^w}{\eta_w} + \epsilon_t^w, \quad \text{with } \epsilon_t^w \sim \text{iid}N(0,\sigma_w^2) \text{ and } \eta^w > 0.$$

Cost minimization implies a demand structure of the form

$$L_t(l) = \left(\frac{W_t(l)}{W_t}\right)^{-\frac{1+\eta_t^w}{\eta_t^w}} L_t,$$

where $W_t(l)$ is the nominal wage rate for type-l labor. The aggregate nominal wage index is then given by

$$W_t = \left(\int_0^1 W_t(l)^{-\frac{1}{\eta_t^w}} dl\right)^{-\eta_t^w}.$$

2.3. **Households.** There is a measure one of households in the economy. The representative household's lifetime utility function writes

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\log\left(C_t - H_t\right) - \frac{1}{1+\kappa} \int_0^1 L_t(l)^{1+\kappa} dl \right),$$

where $\beta \in (0,1)$ is the subjective discount factor and $\kappa \geq 0$ is the inverse Frisch elasticity of labor supply. C_t denotes private consumption, H_t is an external stock of habits that is proportional to the lagged consumption basket:

$$H_t = hC_{t-1}, \quad \text{with } h \in [0,1),$$

and $L_t(l)$ stands for type-l hours worked.

The literature has considered specific mechanisms to help DSGE models generate output multipliers larger than one by boosting private consumption after positive government spending shocks. For instance, Galí, López-Salido, and Vallés (2007) introduce rule-of-thumbs households and Fève, Matheron, and Sahuc (2013) allow for Edgeworth complementarity between private and public consumption. I have tested the empirical relevance of both mechanisms using my estimated model and found them too weak to generate a rise in consumption after a government spending shock. I have thus removed them for simplicity.

The representative household's flow real budget constraint writes

$$C_t + I_t + \frac{B_t}{u_t^b} + \Psi(v_t)K_{t-1} = \frac{R_{t-1}B_{t-1}}{\pi_t} + (1-\tau_t)\left(\int_0^1 \frac{W_t(l)}{P_t} L_t(l)dl + r_t^k v_t K_{t-1} + D_t\right) + Z_t.$$

On the expenditure side, I_t is investment, B_t is real holdings of riskless one-period government bonds, and $\Psi(v_t)K_{t-1}$ is the cost of capital utilization. In the steady state, $v_t = 1$ and the function Ψ is such that $\Psi(1) = 0$. I introduce a parameter $\psi \in [0,1)$ such that $\Psi''(1)/\Psi'(1) = \psi/(1-\psi)$. Also, u_t^b is an exogenous risk premium shock reflecting unmodeled financial frictions. It evolves according to a stationary process:

$$\ln u_t^b = \rho_b \ln u_{t-1}^b + \epsilon_t^b, \quad \text{with } \epsilon_t^b \sim \text{iid}N(0, \sigma_b^2).$$

On the revenue side, $R_{t-1}B_{t-1}/\pi_t$ is real income from bond holdings, $W_t(l)L_t(l)/P_t$ is gross real labor income from type-l labor, $r_t^k v_t K_{t-1}$ is gross real capital income, D_t are dividends, and Z_t is a lump-sum transfer from the government. As in Traum and Yang (2011), I assume that a single income tax rate τ_t applies to both labor and capital income.[2]

Physical capital evolves according to

$$K_t = (1-\delta)K_{t-1} + u_t^i \left[1 - S\left(\frac{I_t}{I_{t-1}}\right)\right] I_t,$$

[2]Allowing for a richer array of tax rates would require augmenting the model with several parameters and the inclusion of additional observables. For the purpose of characterizing tax dynamics after a government spending shock, all rates would follow similar patterns given their common financing role (Leeper, Plante, and Traum, 2010).

where $S(.)$ is an adjustment cost function verifying $S(1) = S'(1) = 0$ and $S''(1) = s$, and u_t^i is a shock to the marginal efficiency of investment evolving according to a stationary process:

$$\ln u_t^i = \rho_i \ln u_{t-1}^i + \epsilon_t^i, \qquad \text{with } \epsilon_t^i \sim \text{iid}N(0, \sigma_i^2).$$

2.3.1. *Wage setting.* Households face Calvo (1983) frictions in nominal wage setting. Each period, they can reoptimize the nominal wage for type-l labor service with probability $1 - \xi_w$. Wages that are not reoptimized are indexed to lagged inflation, according to

$$W_t(l) = \pi_{t-1}^{\iota_w} W_{t-1}(l).$$

When reoptimizing the wage for type-l labor, households solve

$$\max_W \; E_t \sum_{s=0}^{\infty} (\beta \xi_w)^s \left(-\frac{L_{t+s}(l)^{1+\kappa}}{1+\kappa} \right),$$

subject to the demand function for type-l labor, the real budget constraint, and the indexation equation for nominal wages. Letting $W_t^\star$ denote the solution to this maximization program and $w_t^\star = W_t^\star / P_t$ its real counterpart, the aggregate real wage index then evolves according to

$$w_t = \left((1 - \xi_w)(w_t^\star)^{-\frac{1}{\eta_t^w}} + \xi_w \pi_{t-1}^{-\frac{\iota_w}{\eta_t^w}} \pi_t^{\frac{1}{\eta_t^w}} w_{t-1}^{-\frac{1}{\eta_t^w}} \right)^{-\eta_t^w}.$$

2.4. **Public policy.** The monetary authority implements a Taylor-type rule, in which the nominal gross interest rate depends on its lagged value and responds to current inflation and output. Specifically, the monetary rule writes

$$\ln \frac{R_t}{R} = \rho_r \ln \frac{R_{t-1}}{R} + (1 - \rho_r) \left(\eta_\pi \ln \frac{\pi_t}{\pi} + \eta_y \ln \frac{Y_t}{Y} \right) + \ln u_t^m.$$

Variables without time subscript denote steady-state levels, while u_t^m is a disturbance capturing the discretionary component of monetary policy, modeled as a persistent stationary process:

$$\ln u_t^m = \rho_m \ln u_{t-1}^m + \epsilon_t^m, \qquad \text{with } \epsilon_t^m \sim \text{iid}N(0, \sigma_m^2).$$

Each period, the government collects tax revenues and issues one-period nominal riskless bonds $P_t B_t$ to finance its expenditures G_t and Z_t, where G_t represents public consumption. The real government flow budget constraint is thus

$$B_t + \tau_t \left(w_t L_t + r_t^k v_t K_{t-1} + D_t \right) = \frac{R_{t-1} B_{t-1}}{\pi_t} + G_t + Z_t.$$

I follow Leeper, Plante, and Traum (2010) and Coenen, Straub, and Trabandt (2012) in assuming simple policy rules for the three fiscal instruments τ_t, G_t, and Z_t. Each fiscal rule combines four components: an autoregressive term capturing the own persistence of the variable, a response to the lagged debt-to-output ratio reflecting a debt-stabilization motive, a response to the contemporaneous level of activity capturing either automatic stabilization or the effects of loosening the government's budget constraint, and a Gaussian innovation capturing the discretionary component of policy. Letting $s_t = B_t/Y_t$ denote the real debt-to-output ratio, the fiscal rules thus write

$$\widehat{\tau}_t = \rho_\tau \widehat{\tau}_{t-1} + (1-\rho_\tau)(\gamma_\tau \widehat{s}_{t-1} + \kappa_\tau \widehat{Y}_t) + \epsilon_t^\tau,$$
$$\widehat{G}_t = \rho_g \widehat{G}_{t-1} + (1-\rho_g)(\gamma_g \widehat{s}_{t-1} + \kappa_g \widehat{Y}_t) + \epsilon_t^g,$$
$$\widehat{Z}_t = \rho_z \widehat{Z}_{t-1} + (1-\rho_z)(\gamma_z \widehat{s}_{t-1} + \kappa_z \widehat{Y}_t) + \epsilon_t^z.$$

2.4.1. *Remarks on the policy rules.* The above monetary and fiscal rules assume that all policy instruments react to the log deviation between output and its steady state. In contrast, the literature suggests that targeting the output gap (the log deviation between output and its flex-price counterpart) is optimal for both monetary and fiscal authorities (Woodford, 2003; Benigno and Woodford, 2004; Schmitt-Grohé and Uribe, 2006).

Four considerations underlie my modeling choice. First, I use the policy rules as simple statistical representations of the behavior of public authorities, without claiming that they reflect economic optimality. Second, I consider that estimated policy rules are more robust when based on observable variables such as output, rather than on unobservable, model-dependent variables such as the output gap. Third, my specification of policy rules follows preexisting empirical work, allowing for simple comparisons. Fourth, I show in Appendix A that estimation outcomes are reasonably robust to the specification of policy rules and that my baseline choice is associated with the lowest degree of government consumption endogeneity, thereby ensuring that the paper's results are conservative with respect to the design of monetary and fiscal rules.

2.5. **Market clearing and solution method.** Good market clearing requires that

$$Y_t = C_t + I_t + G_t + \Psi(v_t)K_{t-1}.$$

A detailed derivation of the model equilibrium and of the numerical solution method is provided in a technical appendix available upon request. I compute a log-linear approximation to the equilibrium dynamics around the deterministic steady state of the model and use Klein's (2000) approach to solve the resulting rational expectation system.

3. Estimation Results

I estimate the model using Bayesian methods (An and Schorfheide, 2007) and quarterly U.S. series. Because the artificial economy features nine forcing processes, I include nine observables: real consumption, real investment, real government spending, real transfers, real tax revenue, the real wage, hours worked, the nominal interest rate, and the inflation rate. I define government spending as government consumption, to avoid confounding its properties and economic effects with those of government investment. To preserve the coherence of national accounts, I incorporate public investment into the investment series used in estimation and account for it as a transfer from the government to households. This last choice ensures that the model-based debt variable matches its empirical counterpart. Appendix B provides all data sources and describes the linkage to observables.

I remove a linear trend from the logarithms of consumption, investment, government spending, transfers, tax revenue, and the real wage and use the detrended series for estimation. As noted by Leeper, Plante, and Traum (2010), this approach is a simple way to deal with the own trends of fiscal variables, which would otherwise complicate the specification of the model to ensure sustainability.[3] Also, a useful by-product is that I will not need to worry about the trend specification when using the estimated DSGE model to test alternative identification approach in section 4. More importantly, and unlike recent papers focusing on the federal government (Leeper, Plante, and Traum, 2010; Traum and Yang, 2010, 2011), I consider the behavior of an aggregate public sector that also incorporates state and local governments. The rationale behind this choice is the objective of comparability with the empirical literature, where aggregate spending of all governments is typically used.

I consider three estimation samples: 1960Q1-2007Q4, 1960Q1-1978Q4, and 1983Q1-2007Q4. I exclude the 1950s because of the presence of exceptional fiscal shocks caused by the Korean war and I stop the sample in 2007Q4 to avoid nonlinearities due to the binding zero lower bound on the nominal interest rate afterward. I also consider a split sample because several papers, for instance Perotti (2004) and Bouakez, Chihi, and Normandin (2014), document important differences in the U.S. economy's response to fiscal shocks before and after 1979 using SVARs. Allowing for a break should also help avoiding indeterminacy issues due to changing monetary-fiscal policy interactions (Bhattarai, Lee, and Park, 2012). I break the sample according to Galí, López-Salido, and Vallés (2003): the first subsample corresponds to the high inflation period ending with the appointment of Paul Volcker at the Federal Reserve; the second encompasses the so-called 'Great Moderation'; and the 1979Q1-1982Q4 period is excluded on the grounds of its specificity in terms of monetary policy.

3.1. **Prior distributions and calibrated parameters.** As usual, I calibrate some parameters that are difficult to identify from the data. Specifically, I set the discount factor β to 0.99, the depreciation rate δ to 0.025, the capital elasticity of output α to 0.34, and the steady-state markups in the good and labor markets η_p and η_w to 0.10, all standard values. Steady-state gross inflation is normalized to 1. I use averages from the 1960Q1-2007Q4 sample to pin

[3]See Smets and Wouters (2003), Leeper, Plante, and Traum (2010), or Coenen, Straub, and Trabandt (2012) for examples of DSGE models estimated by Bayesian methods on detrended data.

TABLE 1. Selected prior and posterior distributions.

Parameter	Prior distribution			Mode [5%, 95%] of posterior distribution		
	Density	Mean	St.Dev.	1960-2007	1960-1978	1983-2007
Preferences						
κ	$\mathcal{G}$	2.0	0.5	1.90 [1.23, 2.85]	1.82 [1.19, 2.77]	1.84 [1.16, 2.74]
h	$\mathcal{B}$	0.5	0.2	0.93 [0.91, 0.94]	0.90 [0.87, 0.93]	0.91 [0.90, 0.94]
Frictions						
ψ	$\mathcal{B}$	0.6	0.15	0.05 [0.02, 0.11]	0.21 [0.10, 0.43]	0.14 [0.05, 0.24]
s	$\mathcal{G}$	6.0	1.0	8.23 [6.83, 10.10]	5.34 [4.19, 6.84]	7.67 [6.35, 9.80]
ξ_p	$\mathcal{B}$	0.5	0.15	0.91 [0.88, 0.93]	0.86 [0.81, 0.88]	0.90 [0.87, 0.93]
ξ_w	$\mathcal{B}$	0.5	0.15	0.95 [0.92, 0.98]	0.85 [0.78, 0.89]	0.90 [0.83, 0.95]
ι_p	$\mathcal{B}$	0.5	0.15	0.09 [0.04, 0.19]	0.27 [0.18, 0.42]	0.19 [0.08, 0.42]
ι_w	$\mathcal{B}$	0.5	0.15	0.60 [0.46, 0.78]	0.52 [0.36, 0.67]	0.38 [0.18, 0.58]
Fiscal policy						
ρ_τ	$\mathcal{B}$	0.5	0.2	0.97 [0.94, 0.99]	0.85 [0.77, 0.95]	0.90 [0.87, 0.97]
γ_τ	$\mathcal{N}$	0.0	3.0	0.40 [−0.37, 0.56]	0.35 [−0.20, 0.67]	0.24 [−0.13, 1.53]
κ_τ	$\mathcal{N}$	0.0	3.0	0.29 [−1.25, 3.75]	0.09 [−0.38, 1.59]	0.15 [−1.34, 2.68]
ρ_g	$\mathcal{B}$	0.5	0.2	0.95 [0.94, 0.97]	0.90 [0.87, 0.95]	0.94 [0.92, 0.99]
$-\gamma_g$	$\mathcal{N}$	0.0	3.0	0.43 [0.19, 0.70]	0.10 [−0.30, 0.68]	0.17 [−0.16, 1.02]
κ_g	$\mathcal{N}$	0.0	3.0	1.12 [0.49, 1.86]	1.75 [0.77, 2.54]	1.24 [−0.14, 2.90]
ρ_z	$\mathcal{B}$	0.5	0.2	0.81 [0.76, 0.89]	0.75 [0.69, 0.88]	0.75 [0.67, 0.88]
$-\gamma_z$	$\mathcal{N}$	0.0	3.0	0.49 [0.14, 0.99]	1.96 [1.04, 3.21]	0.44 [0.00, 0.99]
$-\kappa_z$	$\mathcal{N}$	0.0	3.0	0.93 [0.00, 2.03]	2.66 [0.79, 4.86]	1.04 [−0.25, 2.47]

down steady-state fiscal ratios: the shares of government spending and transfers in output are $G/Y = 0.16$ and $Z/Y = 0.06$, while the average tax rate is $\tau = 0.22$.

The first columns in Tables 1 and 2 provide the prior distributions for estimated parameters. The priors for standard New Keynesian parameters closely follow the quantitative DSGE literature (Smets and Wouters, 2007; Justiniano, Primiceri, and Tambalotti, 2010). Reflecting Traum and Yang's (2011) findings, I assume that monetary policy is active at the prior mean: the central bank raises the nominal interest rate by more than inflation to ensure price stability. Turning to the debt and output feedback coefficients defining the fiscal rules, I adopt agnostic normal priors with mean 0 and standard deviation 3 to let the data choose both the sign and size of endogenous movements in the average tax rate, government spending, and transfers. All autoregressive coefficients, both in policy rules and in exogenous processes, follow beta priors with mean 0.5 and standard deviation 0.2.

3.2. **Posterior estimates.** The state-space representation of the linearized model allows to evaluate the log-posterior function using the Kalman filter. For each estimation sample, I

TABLE 2. Selected prior and posterior distributions.

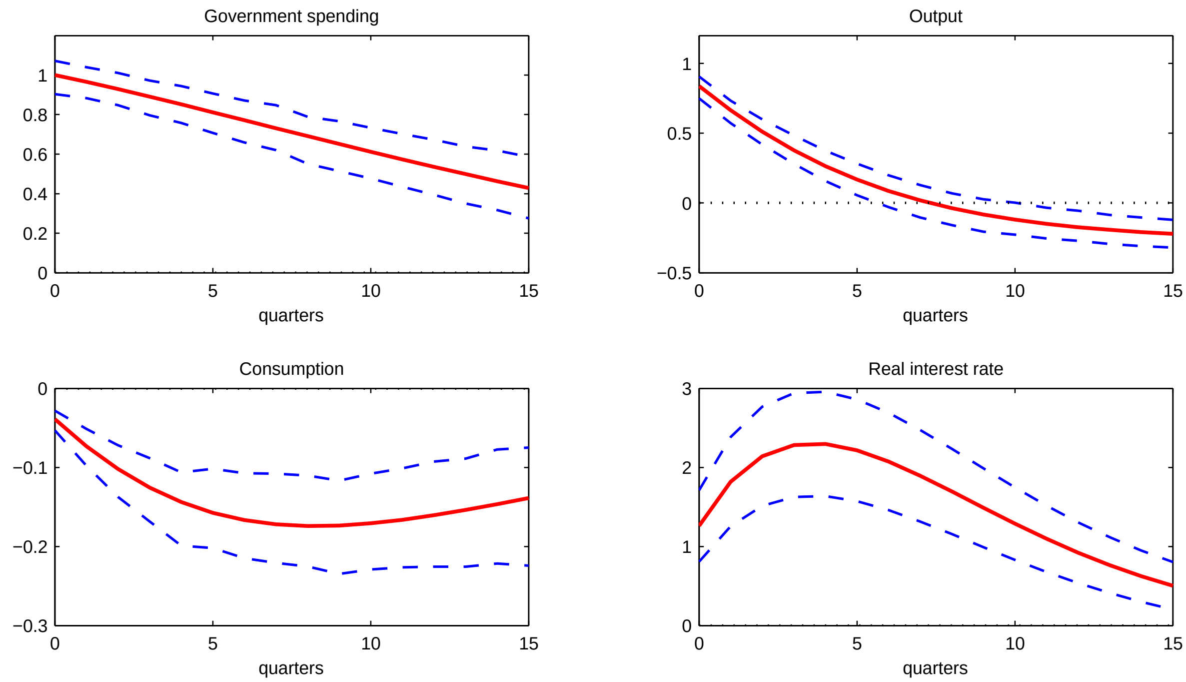
Parameter	Prior distribution			Mode [5%-95%] of posterior distribution		
	Density	Mean	St.Dev.	1960-2007	1960-1978	1983-2007
Monetary policy						
ρ_r	$\mathcal{B}$	0.5	0.2	0.87 [0.81, 0.93]	0.48 [0.39, 0.63]	0.90 [0.86, 0.92]
η_π	$\mathcal{N}$	1.75	0.25	1.38 [1.02, 1.73]	1.50 [1.34, 1.66]	1.77 [1.60, 1.94]
η_y	$\mathcal{N}$	0.2	0.05	0.19 [0.08, 0.36]	0.32 [0.24, 0.43]	0.47 [0.30, 0.64]
Shock processes						
ρ_b	$\mathcal{B}$	0.5	0.2	0.44 [0.31, 0.55]	0.34 [0.18, 0.46]	0.44 [0.28, 0.57]
ρ_a	$\mathcal{B}$	0.5	0.2	0.94 [0.92, 0.96]	0.95 [0.93, 0.96]	0.94 [0.91, 0.96]
ρ_i	$\mathcal{B}$	0.5	0.2	0.26 [0.16, 0.40]	0.23 [0.09, 0.41]	0.47 [0.32, 0.62]
ρ_p	$\mathcal{B}$	0.5	0.2	0.67 [0.54, 0.77]	0.43 [0.23, 0.63]	0.41 [0.11, 0.55]
ρ_w	$\mathcal{B}$	0.5	0.2	0.17 [0.07, 0.27]	0.11 [0.04, 0.25]	0.20 [0.08, 0.35]
ρ_m	$\mathcal{B}$	0.5	0.2	0.26 [0.14, 0.40]	0.98 [0.97, 0.99]	0.76 [0.60, 0.88]
σ_b	$\mathcal{IG}$	3.0	3.0	1.24 [0.86, 1.74]	1.35 [0.88, 1.91]	0.97 [0.63, 1.52]
σ_a	$\mathcal{IG}$	3.0	3.0	2.04 [1.89, 2.25]	2.48 [2.20, 2.88]	1.44 [1.30, 1.64]
σ_i	$\mathcal{IG}$	3.0	3.0	5.69 [4.49, 7.05]	4.44 [3.27, 5.83]	2.78 [2.11, 3.80]
σ_p	$\mathcal{IG}$	3.0	3.0	4.69 [3.60, 6.63]	4.85 [3.35, 7.53]	8.00 [6.10, 11.80]
σ_w	$\mathcal{IG}$	3.0	3.0	3.79 [3.26, 4.36]	3.05 [2.54, 3.58]	4.13 [3.35, 4.95]
σ_m	$\mathcal{IG}$	3.0	3.0	5.02 [4.65, 5.56]	6.38 [5.21, 7.47]	7.52 [6.68, 8.61]
σ_τ	$\mathcal{IG}$	3.0	3.0	1.56 [1.43, 1.70]	1.66 [1.46, 1.91]	1.42 [1.27, 1.61]
σ_g	$\mathcal{IG}$	3.0	3.0	6.92 [6.41, 7.62]	7.95 [7.15, 9.47]	5.50 [4.96, 6.26]
σ_z	$\mathcal{IG}$	3.0	3.0	4.01 [3.72, 4.40]	4.64 [4.11, 5.39]	3.25 [2.95, 3.71]

maximize this function with respect to the estimated parameters and construct the posterior distribution using the random-walk Metropolis-Hastings algorithm with a single Markov chain of one million draws, keeping the last $400,000$ draws for computations. For each chain, I choose a step size ensuring an acceptance rate close to 30% and use standard diagnostic tests to confirm convergence.

The last three columns in Tables 1 and 2 report the means and 90-percent intervals of the posterior distributions for the three estimation samples. Most parameters are well-identified from the data, with the exception of the Frisch elasticity of labor supply. The estimated degree of consumption habits h is substantially above its prior mean. Estimated degrees of nominal rigidities ξ_p and ξ_w are quite high, but introducing strategic complementarities among price setters would mechanically lower them while leaving model dynamics unchanged (Smets and Wouters, 2007).

Turning to policy parameters, there is considerable interest rate smoothing in the full sample, and the central bank's reaction to inflation has become stronger over time. The three fiscal variables are also highly persistent, with autoregressive coefficients estimated close to or above

FIGURE 1. Selected DSGE-estimated responses after a spending shock ϵ^g.



Notes. Solid line: average posterior response. Dashed lines: 90-percent posterior confidence interval. Full sample estimates. The real interest rate is in annualized percentage.

0.80. The estimated policy rules provide significant statistical evidence in favor of endogenous feedback movements in transfers and government spending, albeit not in the tax rate. Namely, transfers respond negatively to both debt and output, reflecting a debt-stabilization motive and the countercyclical character of its most important components like social security payments. More importantly given the focus of the paper, all estimates of γ_g are negative and all estimates of κ_g are positive, implying that government consumption responds negatively to the debt level and positively to contemporaneous output over all samples. This last finding hints that the relaxing of the government budget constraint induced by higher output dominates potential automatic stabilization motives, consistent with Clemens and Miran's (2012) findings for state-level expenditures. Interestingly, the debt elasticity of government appears to increase slightly over time, while the output elasticity instead falls.

Figure 1 illustrates the response of the economy to an exogenous spending shock ϵ_t^g, normalized to raise government expenditures by one unit on impact. All predictions come from the model parametrized with the full-sample estimate, but IRFs are qualitatively and quantitatively similar in the two subsamples.[4] In particular, the impact output multiplier is stable

[4]Output IRFs estimated from the two subsamples are reported in, e.g., Figure 2.

TABLE 3. Model comparisons: Endogenous vs. exogenous government consumption.

Estimation sample	Bayes factor relative to a model with exogenous government consumption
1960-2007	exp(8)
1960-1978	exp(8)
1983-2007	exp(2)

Notes. Bayes factors computed as $p(Y^T \mid \mathcal{M}_1)/p(Y^T \mid \mathcal{M}_2)$, where Y^T is observed data, $\mathcal{M}_1$ is the baseline DSGE model, $\mathcal{M}_2$ is the restricted submodel verifying $\gamma_g = \kappa_g = 0$, and $p(Y^T \mid \mathcal{M}_i)$ is the marginal data density associated with model i. Log marginal data densities computed using the modified harmonic mean estimator.

over time, equal to 0.83 in the full sample and to 0.80 in both subsamples. It is below one because of the crowding out of private consumption and investment, which both fall on impact and over time due to a negative wealth effect and to a persistent rise in the real interest rate induced by the central bank's response to the increase in inflation and output.

3.3. **Endogenous movements in government consumption.** The above estimates suggest that the data support endogenous feedbacks on government consumption. As a more formal measure of the strength of these effects, I use Bayes factors to evaluate the relative fit of the estimated DSGE model compared to one in which exogenous government spending is imposed.

Table 3 reports the Bayes factor associated with each estimation sample. Since the restricted model is the reference, a large statistic signals that the data favor the model with endogenous reactions in government consumption. This is clearly the case in the full sample, for which there is decisive statistical evidence in favor of government spending endogeneity. This is also true for the 1960-1978 period. The statistic for the second subsample is somewhat weaker, but still provides positive evidence in favor of the model with endogenous spending according to Kass and Raftery's (1995) interpretation of Bayes factors. This is an interesting result. Indeed, the estimated coefficients γ_g and κ_g are not different from zero when considered individually. Nevertheless, the Bayes factor shows that the null hypothesis of the joint nullity of γ_g and κ_g is rejected. Therefore, there is significant statistical evidence that government consumption features an endogenous component over all estimation samples. As I show in the next section, the presence of such feedbacks has the potential to induce large bias in estimated output multipliers.

Eventually, it is instructive to look at the respective contributions of the nine structural shock to the forecast error of government consumption, as the fiscal rule implies that all disturbances will affect spending through endogenous effects. In the very short run, the bulk of the forecast error of government consumption is due to the exogenous spending shock in spite of the feedbacks. For instance, more than 97 percent of the one-step-ahead forecast error variance of government consumption originate from ϵ^g in all estimation samples. At longer horizons, other shocks come into play, most notably the monetary disturbance ϵ^m and the tax shock ϵ^τ. The former affects government spending via its large effects on output in this sticky-price economy, while the second propagates to spending via its effect on public deficit and debt. Finally, the contribution of the exogenous government consumption shock to aggregate fluctuations is within the range typically reported in the literature, with a share in the forecast error variance of output close to 3 percent after one quarter, to 2 percent at the one-year horizon, and to 1 percent asymptotically in all samples.

4. Effects of Endogeneity on Estimated Spending Multipliers

The above results suggest that mainstream econometric practices assuming exogenous government consumption may be at odds with the data. This raises the question of the effects of omitting to control for endogeneity when identifying output multipliers in general equilibrium. In this section, I use simple Monte Carlo experiments to quantify these effects and assess the performances of three estimation approaches: a DSGE model with exogenous spending, a VAR model identified with Blanchard and Perotti's (2002) exogeneity restriction, and a VAR identified with Mountford and Uhlig's (2009) sign restrictions. The two first methods assume away spending endogeneity, while the last one tries to control for it.

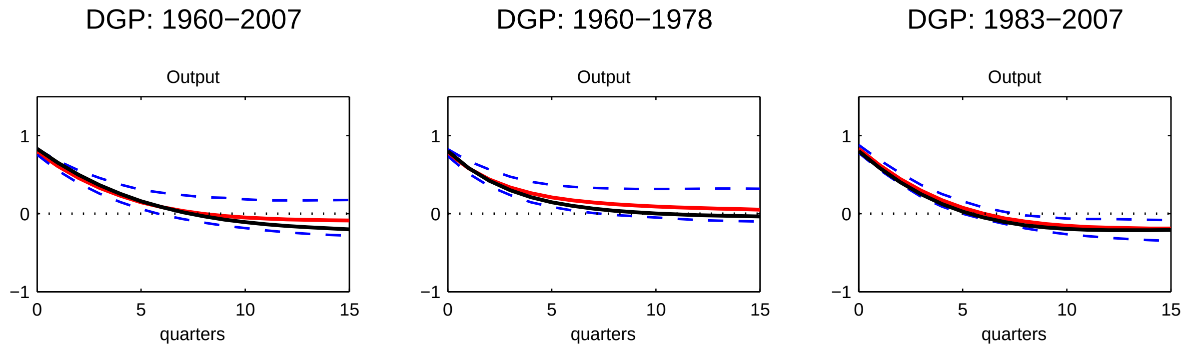
I design the experiments as follows. Taking the estimated DSGE model as data generating process (DGP), I simulate 1,000 artificial time series and identify spending shocks and output multipliers using the three approaches under test. I report results under three forms: output IRFs corresponding to an exogenous spending shock increasing spending by 1 dollar on impact, present-value multipliers at horizons 0 and 4, and average correlations between the true spending shock in the DGP, ϵ_t^g, and that identified by the econometrician, denoted $\widehat{\epsilon}_t^g$. I define present-value multipliers as

$$PVM(k) = \frac{\sum_{j=0}^{k} R^{-j} dY_{t+j}}{\sum_{j=0}^{k} R^{-j} dG_{t+j}},$$

where R is the steady-state gross nominal interest rate and dY_{t+j} and dG_{t+j} denote respectively the responses of output and government consumption j periods after the impulse. Thus, $PVM(0)$ corresponds to a standard impact multiplier, while $PVM(4)$ accounts for the full dynamics of the economy one year after the shock.

I work with three parameterizations of the DGP, corresponding respectively to the point estimates obtained for the 1960-2007, 1960-1978, and 1983-2007 periods. For each design, I include as many observations as in the original sample used to estimate the DSGE model, that is 192, 76, and 100 respectively. I always discard a burn-in period of 500 observations. Full-sample experiments closely correspond to empirical papers abstracting from breaks, while

FIGURE 2. Estimation of output IRFs: DSGE model with exogenous spending.



Notes. Solid black line: IRF to the spending shock in the DGP (DSGE model with endogenous policy). Solid red line: average IRF estimated using a DSGE model with exogenous policy. Dashed blue lines: empirical 90% confidence interval.

subsample experiments allow for a sensitivity analysis given the changing patterns of spending endogeneity over time. Figures 2-5 and Table 4 display the results.

From an econometric perspective, the experiments have an indirect inference flavor since all tested models are misspecified (Gourieroux, Monfort, and Renault, 1993). Unlike indirect inference however, this misspecification is not taken into account when computing multipliers. Formally, the experiments thus quantify estimation errors arising with misspecified model omitting endogenous feedback effects on government consumption.

4.1. **DSGE model with exogenous spending.** I first test the performance of a restricted version of the DSGE model in which government consumption exogeneity is assumed. For each artificial dataset, I reestimate the model imposing $\kappa_g = \gamma_g = 0$, using the same set of observables as before and the same prior distributions as in Tables 1-2. I compute IRFs and multipliers at the posterior mode.

The three panels in Figure 2 compare the average estimated output IRF with that implied by the true DGP for the three parameterizations. Overall, the performance of the misspecified DSGE model seems satisfactory, as the true output response always lies within the tight confidence bands around the average estimated IRF. This is true with all three DGPs, irrespective of whether they feature significant feedbacks from both debt and output on government consumption (first panel, 1960-2007 DGP), a significant feedback from contemporaneous output only (central panel, 1960-1978 DGP), or no independently significant feedback (last panel, 1983-2007 DGP). The statistics in Table 4 confirm that point estimates of multipliers are accurately estimated, with bias of less than five percent on impact and less than ten percent at the one-year horizon. Also, sampling uncertainty is moderate, allowing to perform precise inference on the size of true multipliers. Eventually, in all designs the identified spending shocks are strongly linked with the true ones, with average correlations above 0.9.

One interesting outcome of the experiment is the fact that impact multipliers are slightly underestimated with DGPs corresponding to the full 1960-2007 sample and to the first 1960-1978 subsample. Given the procyclical response of government consumption to contemporaneous output, the analytical argument in Fève, Matheron, and Sahuc (2013) would instead imply that output multipliers should be overestimated when neglecting the endogenous component. It is in fact the presence of the debt feedback, not considered by Fève, Matheron, and Sahuc, which explains this discrepancy. In the true DGP, a positive, exogenous innovation to government consumption raises the debt-to-DGP ratio on impact, which will have a negative effect on the future path of spending. Neglecting this feedback effect yields one to underestimate the persistence of government consumption, which is key for the size of the economy's response.

4.2. **VARs identified with exogeneity restrictions.** I now turn to VAR models identified with the exogeneity restrictions pushed forward by Blanchard and Perotti (2002). I consider the generic structural model

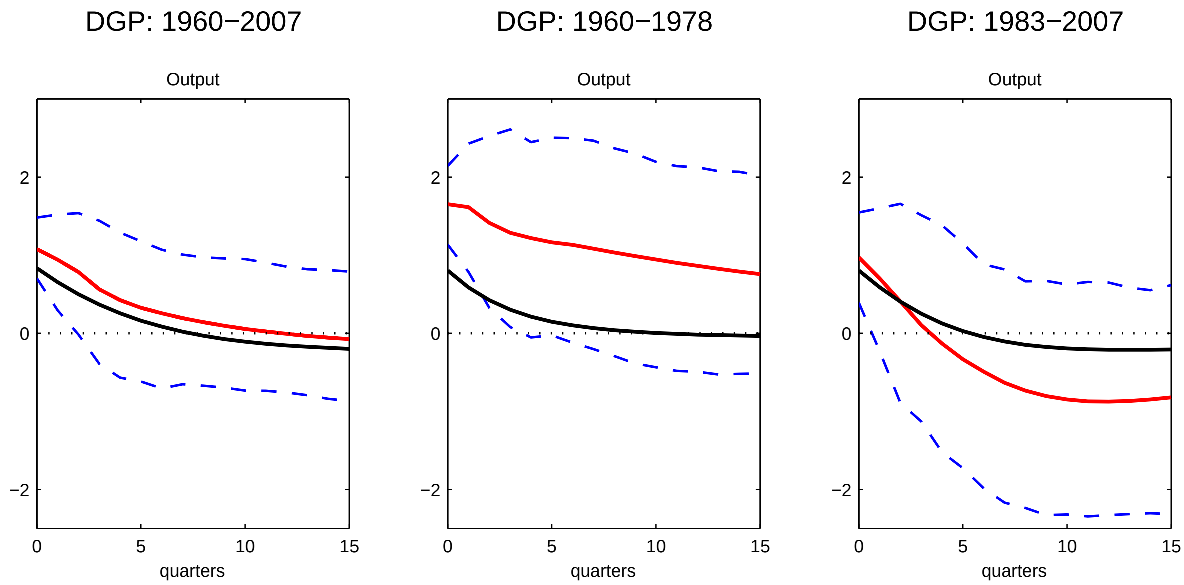
$$\mathbf{A}\begin{bmatrix} G_t \\ Y_t \\ \mathbf{X}_t \end{bmatrix} = \mathbf{B}(L)\begin{bmatrix} G_{t-1} \\ Y_{t-1} \\ \mathbf{X}_{t-1} \end{bmatrix} + \boldsymbol{\eta}_t, \tag{1}$$

where G_t and Y_t denote government consumption and output, $\mathbf{X}_t$ is a vector of additional observables, $\mathbf{B}(L)$ is a matrix polynomial of order p, and $\boldsymbol{\eta}_t$ is a vector of 'structural' shocks. Variables are expressed as deviations from their steady state, so I omit the constant. I estimate the reduced-form version of Eq. (1) on artificial data and identify the spending shock by imposing that government consumption cannot react within a period to structural shocks other than its own innovations. Formally, this amounts to imposing that the first row of $\mathbf{A}$ contains a one in its first entry and zeros elsewhere.

In practice, I test several choices of $\mathbf{X}_t$. The benchmark specification is the baseline 3-variable system from Blanchard and Perotti (2002) that includes tax revenue as only additional observable. Because the DGP features a debt feedback, I also study a 4-variable system including both taxes and the debt-to-output ratio. Eventually, I test a larger system in which $\mathbf{X}_t$ includes, on top of taxes and debt, consumption, hours worked, the nominal interest rate, and inflation. In all cases, I set $p = 4$ after checking that the results are not sensitive to further increase in the lag length.

Figures 3 and 4 display the average output IRFs identified from the 3-variable and 4-variable VARs. Results for the larger system are comparable to those for the 4-variable VAR, so I omit them to save space. It is striking that in all designs, exogeneity restrictions overestimate

Figure 3. Estimation of output IRFs: 3-variable VAR identified with exogeneity restrictions.

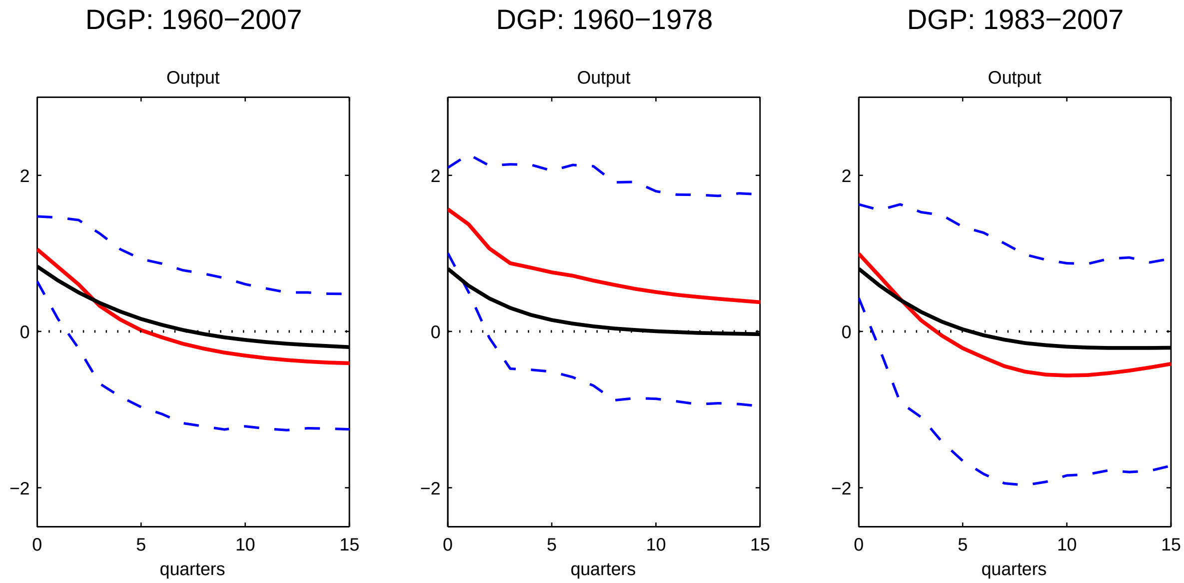


Notes. Solid black line: IRF to the spending shock in the DGP. Solid red line: average IRF estimated using a VAR identified via exogeneity restrictions. Dashed blue lines: empirical 66% confidence interval. The VAR includes government spending, output, and tax revenue as observables.

the output response on impact. The resulting upward bias in output multipliers is quantitatively large, ranging from 20 percent when the DGP matches the properties of the 1983-2007 subsample to almost 110 percent when the DGP corresponds to the first subsample. Another remarkable result is the change in the shape of estimated IRFs over time: in the first subperiod, the average VAR-estimated output response to a government consumption shock is strongly persistent and stays above zero for several years, whereas it quickly decays and falls below zero after one year in the second subsample. Such patterns are in line with subsample VAR estimates reported in the literature (see, e.g., Perotti, 2004), confirming that the DSGE model used as DGP captures important properties of the data. However, the experiment emphasizes that differences in endogeneity patterns rather than in the underlying theoretical output response are responsible for these changing estimates.

To gain intuition on the sources of the identification problem, it is instructive to compare the results across Figures 3 and 4. In Figure 3, the estimated VAR does not include debt among observables. Since the lagged values of spending, output, and tax revenue provide an incomplete signal for the budget balance because of the omission of transfers, it is impossible for this VAR to correctly capture the debt feedback. On the other hand, the VAR estimated in Figure 4 includes an observation of debt and solves this omitted variable problem, but it still

FIGURE 4. Estimation of output IRFs: 4-variable VAR identified with exogeneity restrictions.



Notes. Solid black line: IRF to the spending shock in the DGP. Solid red line: average IRF estimated using a VAR identified via exogeneity restrictions. Dashed blue lines: empirical 68% confidence interval. The VAR includes government spending, output, tax revenue, and the debt-to-GDP ratio as observables.

yields severely biased output IRFs and spending multipliers. The only effect of incorporating debt in the system is to lower the persistence of the estimated output response, but it does not improve estimates of the impact multiplier. The econometric issue therefore lies in the inability of exogeneity restrictions to identify the true spending shock in presence of a contemporaneous output feedback. As discussed in Fève, Matheron, and Sahuc (2013), the underlying intuition is straightforward. On the one hand, exogeneity restrictions implicitly attribute all the conditional correlation between output and government consumption to the multiplier. On the other, the presence of a fiscal rule means that part of this correlation reflects endogenous mechanisms. The combination of these two arguments implies that exogeneity restrictions will overestimate the output multiplier in presence of a procyclical government spending rule.

Another important property uncovered by the experiments is the large amount of sampling uncertainty in output IRFs and multipliers derived from the VARs. The confidence bands displayed in Figures 3 and 4 have a 68-percent coverage, yet they are much larger than the 90-percent confidence bands for DSGE-based estimates shown in Figure 2.[5] As can be seen from Table 4, the empirical standard deviations for estimated multipliers are also often more than ten times larger with VARs than with DSGE models. At one level, this pattern is not

[5]I have chosen to display 68-percent confidence bands for VAR experiments to avoid overloading the figures.

TABLE 4. Estimation of Present-Value Multipliers on Artificial Data.

	Data Generating Process		
	1960-2013	1960-1978	1983-2007
Theoretical multipliers in the DGP			
$PVM(0)$	0.833	0.803	0.803
$PVM(4)$	0.627	0.600	0.553
DSGE model with exogenous government spending			
$PVM(0)$	0.794 (0.017)	0.780 (0.020)	0.834 (0.022)
$PVM(4)$	0.556 (0.032)	0.564 (0.038)	0.608 (0.040)
$\text{corr}(\epsilon^g, \widehat{\epsilon}^g)$	0.947 (0.020)	0.903 (0.035)	0.955 (0.020)
VAR identified with exogeneity restrictions (3-variable system)			
$PVM(0)$	1.077 (0.377)	1.653 (0.511)	0.972 (0.562)
$PVM(4)$	0.839 (0.582)	1.596 (0.852)	0.597 (1.111)
$\text{corr}(\epsilon^g, \widehat{\epsilon}^g)$	0.944 (0.017)	0.899 (0.032)	0.910 (0.030)
VAR identified with sign restrictions			
$PVM(0)$	1.724 (1.179)	2.356 (1.206)	1.910 (1.681)
$PVM(4)$	1.605 (1.334)	2.295 (1.155)	1.910 (1.598)
$\text{corr}(\epsilon^g, \widehat{\epsilon}^g)$	0.220 (0.051)	0.200 (0.050)	0.237 (0.053)

Notes. Estimated multipliers are sample averages over the replications. Standard deviations in parenthesis.

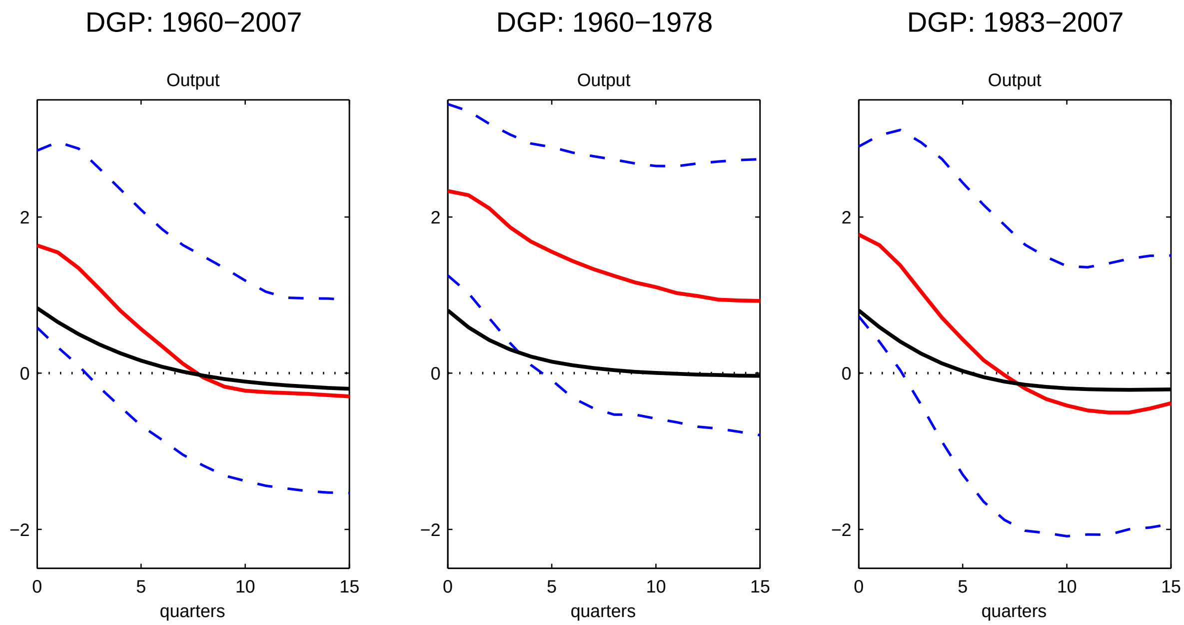
surprising, as even the 3-variable VAR has more coefficients that the DSGE model. However, it clearly calls into question the robustness of multipliers estimated from short subsamples using VARs, such as those discussed in Perotti (2004) or Bouakez, Chihi, and Normandin (2014).

To sum up, the experiments thus suggest that multipliers derived from VARs identified via exogeneity restrictions are less robust than their DSGE counterparts: they are more severely affected by the presence of an endogenous component in government spending and are much more volatile.

4.3. **VARs identified with sign restrictions.** The last econometric model I test is a VAR identified with sign restrictions. Mountford and Uhlig (2009) push forward this approach because of its apparent ability to filter out automatic movements in fiscal variables over the business cycle. It is thus of particular interest to check its empirical performance against DGPs estimated from the data and allowing for endogenous shifts in fiscal policy.

Practically, I start from the structural VAR (1), in which $\mathbf{X}_t$ includes tax revenue, the debt-to-output ratio, consumption, hours worked, the nominal interest rate, and inflation. Using a large number of observables is helpful to strengthen identification via sign restrictions. I identify only two disturbances: a business cycle shock, defined as a shock that jointly moves output, tax revenue, consumption, and hours worked for four periods, and a government spending shock,

Figure 5. Estimation of output IRFs: VAR identified with sign restrictions.



Notes. Solid black line: IRF to the spending shock in the DGP. Solid red line: average IRF estimated using a VAR identified by sign restrictions. Dashed blue lines: empirical 68% confidence interval.

defined as a shock orthogonal to the business cycle shock and that increases spending for four periods.[6] I perform identification using Uhlig's (2005) 'pure-sign-restriction' approach: for each artificial dataset, I search for factorizations of the residual covariance matrix that verify the sign restrictions using random draws, keeping 200 valid candidates per dataset. The output IRF for this particular realization of the DGP is then the average response over the candidates.

Results are plotted in Figure 5. Remarkably, average output IRFs identified by sign restrictions are more biased that those obtained by exogeneity restrictions, even though the very objective of sign restrictions is to control for feedback effects. The impact response of output to the government consumption shock is largely overestimated in all designs, and this is also true of estimated multipliers. On impact, upward bias in multipliers range from 100 to 200 percent, and can reach up to 300 percent at the one-year horizon. Another sign of the poor performance of sign restrictions is provided by the correlations between the VAR-identified spending shocks and the true government consumption disturbances, which are much lower than with the two other econometric approaches. Eventually, sampling uncertainty is large and estimation outcomes are very noisy.

[6]For computational reasons, I do not identify a monetary shock. Both Caldara and Kamps (2012) and preliminary work for this paper confirm that this omission has little effect on the outcome of the sign-restriction approach.

Because the estimated VAR includes the debt-to-output ratio, the debt feedback plays no important role here. Again, the issue lies with inability of sign restrictions to deal with the output feedback. The experiments thus provide little support for the claim that multipliers identified by sign restrictions are more robust to endogenous movements in fiscal variables. This is a deceptive finding, especially given the computational cost of practically implementing the sign restrictions.

5. Conclusion

The objectives of this paper are twofold. First, I use an estimated DSGE model with endogenous fiscal policy to quantify the feedback effects affecting government consumption in the U.S. economy. I find significant statistical evidence in favor of such endogenous patterns over both a long sample and two subperiods, even though the strength of the feedbacks has somewhat decreased over time. Second, I use the estimated model as laboratory to test spending multipliers derived from a DSGE model with exogenous policy, VARs identified with exogeneity restrictions, and VARs identified with sign restrictions. The experiments suggest that multipliers are quite accurately estimated by DSGE models but may be severely overestimated and noisy when derived from SVARs. Importantly, I find that sign restrictions perform worse that simpler exogeneity restrictions, even though they are supposed to deal with endogenous movements in fiscal policy.

My results suggest some interesting research avenues for future work. First, it would be important to provide some microfoundation for the feedback policy effects and Clemens and Miran's (2012) analysis of the role of balanced-budget rules for local governments provides an interesting starting point. Second, it would be instructive to disaggregate government consumption into federal and local components, or into different types of expenditures, to refine the empirical analysis of the feedbacks. Eventually, the methodology implemented in this paper provides a natural tool to evaluate tax output multipliers, whose identification certainly suffers from similar endogeneity issues.

References

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172. [9]

Benigno, P., and M. Woodford (2004): "Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach," in *NBER Macroeconomics Annual 2003, Volume 18*, NBER Chapters, pp. 271–364. National Bureau of Economic Research, Inc. [9]

Bhattarai, S., J. W. Lee, and W. Y. Park (2012): "Monetary-Fiscal Policy Interactions and Indeterminacy in Postwar US Data," *American Economic Review*, 102(3), 173–78. [10]

Blanchard, O., and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spendings and Taxes on Output," *Quarterly Journal of Economics*, 117, 1329–1368. [3, 15, 17]

Bouakez, H., F. Chihi, and M. Normandin (2014): "Measuring the Effects of Fiscal Policy," *Journal of Economic Dynamics and Control*, 47, 123–151. [4, 10, 20]

Caldara, D., and C. Kamps (2012): "The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers," Finance and Economics Discussion Series 2012-20, Board of Governors of the Federal Reserve System (U.S.). [21]

Calvo, G. A. (1983): "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12(3), 383–398. [6, 8]

Chahrour, R., S. Schmitt-Grohé, and M. Uribe (2012): "A Model-Based Evaluation of the Debate on the Size of the Tax Multiplier," *American Economic Journal: Economic Policy*, 4, 28–45. [4]

Christiano, L. J., M. Eichenbaum, and S. Rebelo (2011): "When Is the Government Spending Multiplier Large?," *Journal of Political Economy*, 119(1), 78 – 121. [3]

Clemens, J., and S. Miran (2012): "Fiscal Policy Multipliers on Subnational Government Spending," *American Economic Journal: Economic Policy*, 4(2), 46–68. [3, 4, 13, 22]

Coenen, G., R. Straub, and M. Trabandt (2012): "Fiscal Policy and the Great Recession in the Euro Area," *American Economic Review*, 102(3), 71–76. [8, 10]

Cogan, J. F., T. Cwik, J. B. Taylor, and V. Wieland (2010): "New Keynesian versus Old Keynesian Government Spending Multipliers," *Journal of Economic Dynamics and Control*, 34(3), 281–295. [3]

Cúrdia, V., A. Ferrero, G.-C. Ng, and A. Tambalotti (2015): "Has U.S. Monetary Policy Tracked the Efficient Interest Rate?," *Journal of Monetary Economics*, 70(C), 72–83. [25]

Drautzburg, T., and H. Uhlig (2011): "Fiscal Stimulus and Distortionary Taxation," NBER Working Papers 17111, National Bureau of Economic Research, Inc. [3]

Fatás, A., and I. Mihov (2001): "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," CEPR Discussion Papers 2760, C.E.P.R. Discussion Papers. [3]

Fève, P., J. Matheron, and J.-G. Sahuc (2013): "A Pitfall with DSGE-Based, Estimated, Government Spending Multipliers," *American Economic Journal: Macroeconomics*, 5, 141–178. [3, 7, 17, 19]

Galí, J., J. D. López-Salido, and J. Vallés (2003): "Technology Shocks and Monetary Policy: Assessing the Fed's Performance," *Journal of Monetary Economics*, 50(4), 723–743.

[10]

——— (2007): "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association*, 5(1), 227–270. [3, 7]

Gourieroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, 85–118. [16]

Jones, J. B. (2002): "Has Fiscal Policy Helped Stabilize the Postwar US Economy?," *Journal of Monetary Economics*, 49, 709–746. [3]

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010): "Investment Shocks and Business Cycles," *Journal of Monetary Economics*, 57, 132–145. [11]

Kass, R. E., and A. E. Raftery (1995): "Bayes Factors," *Journal of the American Statistical Association*, 90(430), 773–795. [14]

Klein, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectation Model," *Journal of Economic Dynamics and Control*, 24, 1405–1423. [9]

Leeper, E. M., M. Plante, and N. Traum (2010): "Dynamics of Fiscal Financing in the United States," *Journal of Econometrics*, 156, 304–321. [3, 7, 8, 10]

Leeper, E. M., N. Traum, and T. B. Walker (2015): "Clearing Up the Fiscal Multiplier Morass," manuscript. [25]

McGrattan, E. R. (1994): "The Macroeconomic Effects of Distorsionary Taxation," *Journal of Monetary Economics*, 33, 573–601. [3]

Mountford, A., and H. Uhlig (2009): "What Are the Effects of Fiscal Policy Shocks?," *Journal of Applied Econometrics*, 24, 960–992. [4, 15, 20]

Perotti, R. (2004): "Estimating the Effects of Fiscal Policy in OECD Countries," Working Papers 276, IGIER, Bocconi University. [10, 18, 20]

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's all in the Timing," *The Quarterly Journal of Economics*, 126(1), 1–50. [4]

Schmitt-Grohé, S., and M. Uribe (2006): "Optimal Fiscal and Monetary Policy in a Medium-Scale Macroeconomic Model," in *NBER Macroeconomics Annual 2005, Volume 20*, NBER Chapters, pp. 383–462. National Bureau of Economic Research, Inc. [9]

Smets, F., and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123–1175. [10]

——— (2007): "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606. [3, 11, 12]

Traum, N., and S.-C. S. Yang (2010): "When Does Government Debt Crowd Out Investment?," CAEPR Working Papers 2010-006, Center for Applied Economics and Policy Research, Economics Department, Indiana University Bloomington. [10]

——— (2011): "Monetary and Fiscal Policy Interactions in the Post-War U.S.," *European Economic Review*, 55(1), 140–164. [5, 7, 10, 11]

Uhlig, H. (2005): "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, 52(2), 381–419. [21]

——— (2010): "Some Fiscal Calculus," *American Economic Review*, 100(2), 30–34. [3]

Woodford, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press, Princeton, New Jersey, USA. [9]

Appendix A. Robustness

This appendix discusses the robustness of the estimation results from section 3 to the specification of fiscal and monetary policy rules.

First, I generalize the government spending policy rule to

$$\widehat{G}_t = \rho_{g1}\widehat{G}_{t-1} + \rho_{g2}\widehat{G}_{t-2} + (1-\rho_{g1}-\rho_{g2})(\gamma_g \widehat{s}_{t-1} + \kappa_{g1}\widehat{Y}_t + \kappa_{g2}\widehat{Y}_{t-1}) + \epsilon_t^g. \tag{F1}$$

Compared to the benchmark, this specification allows for richer dynamics thanks to the AR(2) structure and incorporates feedbacks from both contemporaneous and lagged output. Under parameter restrictions, it also collapses to an AR(1) rule specified in terms of the spending-to-output ratio. Second, motivated by Leeper, Traum, and Walker's (2015) contention that monetary policy plays a key role in shaping the size of output multipliers, I consider alternative rules for the nominal interest rate:

$$\widehat{R}_t = \rho_r \widehat{R}_{t-1} + (1-\rho_r)[\eta_\pi \widehat{\pi}_t + \eta_y(\widehat{Y}_t - \widehat{Y}_{t-1})] + \epsilon_t^m, \tag{M1}$$

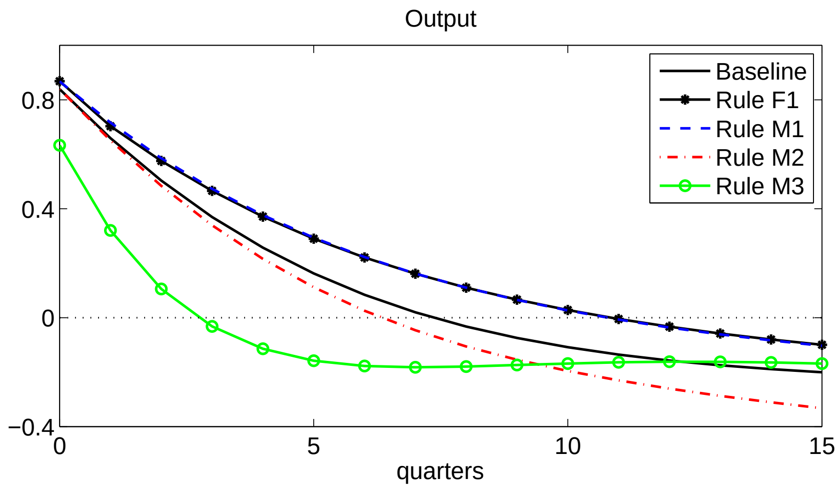
$$\widehat{R}_t = \rho_r \widehat{R}_{t-1} + (1-\rho_r)[\eta_\pi \widehat{\pi}_t + \eta_y(\widehat{Y}_t - \widehat{Y}_t^\star)] + \epsilon_t^m, \tag{M2}$$

$$\widehat{R}_t = \rho_r \widehat{R}_{t-1} + (1-\rho_r)(\widehat{r}_t^\star + \eta_\pi \widehat{\pi}_t) + \epsilon_t^m. \tag{M3}$$

Rule (M1) assumes that the central bank sets the nominal interest rate in reaction to output growth rather than output. According to rule (M2), the nominal interest rate responds instead to the output gap, with $Y_t^\star$ representing the efficient output level in an economy without nominal rigidity. Eventually, rule (M3) follows Cúrdia, Ferrero, Ng, and Tambalotti (2015) by assuming that the monetary authority tracks the efficient real interest rate $r_t^\star$ that would obtain in an economy without nominal friction.

Figure 6 reports the output responses to an exogenous government consumption shock that obtain when estimating each of the above specifications from the full 1960-2007 sample. To ease comparison, I also report the baseline output IRF estimated from the benchmark model. Rules (F1), (M1), or (M2) yield impact output multipliers very close to the baseline specification, but the economy's response to the spending shock is slightly more persistent using rules (F1) and (M1). On the other hand, rule (M3) is associated with a smaller impact multiplier and a shorter-lived positive output response. Yet, it is also the least preferred specification for the monetary policy rule according to Bayes factors.

FIGURE 6. DSGE-estimated output responses after a spending shock: Robustness.



Notes. Solid line: average posterior responses. Full-sample estimates.

Eventually, I emphasize that rules (F1), (M1), (M2), and (M3) are all associated with *higher* degrees of government spending endogeneity compared to the benchmark choice.[7] Specifically, the baseline specification is associated with the smallest point estimates of the debt and output elasticity of government consumption. Given that this contemporaneous feedback is at the heart of the identification issues studied in section 4, the paper's results may be viewed as conservative with respect to the design of policy rules.

APPENDIX B. DATA CONSTRUCTION

This appendix describes sources and data construction for the nine time series used for estimation. I convert nominal series to real values by dividing them by the implicit deflator for personal consumption expenditures (NIPA Table 1.1.4, line 2). Real series are expressed in per-capita terms using the civilian non-institutional population over 16 (BLS, LNU00000000Q).

Consumption. Private consumption, C_t, is defined as consumption expenditures on non-durable goods and services (BEA, NIPA table 1.1.5, lines 5 and 6).

Investment. Gross investment, I_t, is defined as the sum of consumption expenditures on durable goods (BEA, NIPA table 1.1.5, line 4), gross private domestic investment (BEA, NIPA table 1.1.5, line 7), and gross government investment (BEA, NIPA table 3.1, line 36).

Government spending. Government spending, G_t, is defined as government consumption expenditures (BEA, NIPA table 3.1, line 18).

[7] Detailed estimation results for all specifications are available upon request.

Tax revenue. Tax revenue, T_t, is defined as the sum of personal current taxes (BEA, NIPA table 3.1, line 3), half of taxes on production and imports (BEA, NIPA table 3.1, line 4), taxes on corporate income (BEA, NIPA table 3.1, line 5), and contributions for government social insurance (BEA, NIPA table 3.1, line 7). I do not fully include taxes on production and imports because they partly include excise taxes akin to a distortionary tax on consumption. NIPA table 3.5 suggests that a half-half representation provides a good approximation to the distribution.

Transfers. I define transfers, Z_t, as the residual in the government budget constraint to ensure that the model-based debt variable matches its empirical counterpart. Therefore, transfers are defined as the sum of net transfer payments, net capital transfer payments, subsidies (BEA, NIPA table 3.1, line 27), and other government expenditures, minus half of taxes on production and imports, taxes from the rest of the world (BEA, NIPA table 3.1, line 6), income receipts on assets (BEA, NIPA table 3.1, line 8), and current surplus of government enterprises (BEA, NIPA table 3.1, line 16). Net transfer payments are defined as current transfer payments (BEA, NIPA table 3.1, line 19) minus current transfer receipts (BEA, NIPA table 3.1, line 13), while net capital transfer payments are defined as capital transfer payments (BEA, NIPA table 3.1, line 37) minus capital transfer receipts (BEA, NIPA table 3.1, line 33). Other government expenditures are defined as the sum of gross government investment (BEA, NIPA table 3.1, line 36) and government purchases of nonproduced assets (BEA, NIPA table 3.1, line 38), minus government consumption of fixed capital (BEA, NIPA table 3.1, line 39).

Hours worked. Hours worked, L_t, are defined as $H_t \times N_t$, where H_t denotes average nonfarm business weekly hours duration (BLS, PRS85006023) and N_t denotes civilian employment (BLS, CE16OV).

Wage rate. The wage rate, W_t, is defined as the index for hourly compensation in the nonfarm business sector (BLS, PRS85006103).

Inflation. The gross inflation rate, π_t, is defined as the growth rate of the implicit deflator for personal consumption expenditures (BEA, NIPA table 1.1.4, line 2).

Nominal interest rate. The nominal interest rate, R_t, is constructed from the quarterly average of the effective federal funds rate (FRED database).