Propagation Mechanisms for Government Spending Shocks: A Bayesian Comparison*

Anna Kormilitsina† Southern Methodist University

Sarah Zubairy ‡ Texas A&M University

January 1, 2018

Abstract

The inability of a simple real business cycle model to predict a rise in consumption in response to increased government expenditures, observed in many empirical studies, has stimulated the development of alternative theories of government spending shocks. Using the Bayesian approach, we evaluate the quantitative performance of five extant models, and find that neither of the considered transmission mechanisms for government spending helps improve the fit of the baseline model. Moreover, we find that consumption decreases in all estimated models in response to a rise in government spending.

JEL Classification: C11, E32, E62

Keywords: government spending shock, Bayesian comparison, transmission mechanism, deep habits, rule of thumb consumers

*We would like to thank Chetan Dave, Stephanie Schmitt-Grohe and conference participants at 2012 Midwest Macro Meetings, 2013 Canadian Economic Association Meetings and 2013 SNDE conference, as well as two anonymous referees for comments and suggestions. All errors are our own.

†E-mail: annak@smu.edu

‡E-mail: szubairy@tamu.edu

1 Introduction

Recently, there has been a rising interest in modeling government spending and its effects on the economy. This growing research has resulted in a variety of models of government spending shocks, however the question remains as to which model is most appropriate for empirical analysis. In this paper, we use a medium-scale dynamic stochastic general equilibrium (DSGE) framework to quantitatively investigate several propagation mechanisms for government spending shocks proposed in the literature.

All the models we include in this investigation were developed in an attempt to resolve the inconsistency between empirical and theoretical literature predictions about the co-movement between public and private expenditures conditional on a government spending shock. The response of consumption to a government spending shock is subject to a lively debate, and is of great importance in studying the stimulative effects of increased government spending. While some empirical studies find that a government spending increase will boost private consumption (see Blanchard and Perotti (2002), Fatas and Mihov (2001), Mountford and Uhlig (2009) and Fisher and Peters (2010)), traditional RBC models fail to generate this positive correlation between private and public expenditures.[1] The main reason for this is that an increase in government spending generates a dominating negative wealth effect on consumers, which inevitably leads to a fall in private consumption.

Different modifications to a standard model have been proposed in the literature. For example, Ravn, Schmitt-Grohé, and Uribe (2006) argue that positive response of consumption to a government spending shock can be achieved if firm markups of prices over marginal costs are counter-cyclical with the economic activity. In this situation, since rising government spending results in an expansion of aggregate demand, the markups fall and as a consequence, the labor demand rises. With the sufficient expansion of the labor demand and hours in equilibrium, wages may consequently go up to ensure a rise in consumption. Countercyclical movements in markups can generally be achieved by introducing price stickiness. However, Linnemann and Schabert (2003) demonstrate that price stickiness alone is not sufficient to predict a rise in consumption in response to increasing government expenditures. Ravn, Schmitt-Grohé, and Uribe (2006) use the notion of "deep habits" in preferences for consumption to generate endogenous countercyclical markups. Deep habit formation implies that consumers form habits at individual varieties of goods, rather than at the aggregate level, as is the case in more standard models of "superficial" habit formation. Ravn, Schmitt-Grohé, and Uribe (2006) show that the deep habits mechanism can poten-

[1] The response of consumption to a government spending shock is not uncontroversial, with the empirical literature predicting positive, insignificant and negative responses to government spending shock. See Perotti (2008) and Ramey (2011) for a discussion.

tially generate movements in markups large enough to guarantee a rise in consumption even in the absence of price stickiness.

An alternative way to model positive correlation between public and private consumption is offered by Galí, Lopez-Salido, and Vallés (2007). The authors introduce households who do not make optimizing decisions, and may therefore increase consumption in response to a rise in government spending. Following the so-called "rule-of-thumb", these households consume their entire disposable income in each period. If an increase in consumption of the rule-of-thumb households is large enough, the aggregate consumption may increase after rising government spending.

Besides deep habit formation and rule-of-thumb households, other modifications of a standard RBC framework have been used to resolve the problem of co-movement between private and public consumption. Firstly, Linnemann and Schabert (2004) and Bouakez and Rebei (2007) consider an environment where the household directly benefits from government spending through increased utility. They show that if the elasticity of substitution between public and private spending is sufficiently low, then an increase in government spending raises the marginal utility of consumption, making private consumption more attractive for households. If this effect dominates the negative wealth effect of public spending, the positive correlation of private and public consumption may be observed in response to a public spending shock. Ganelli and Tervala (2009) arrive at the same conclusion in a model where public and private consumption are complements.

Secondly, Baxter and King (1993), Ambler and Paquet (1996), and Linnemann and Schabert (2006) model government spending as enhancing productivity of firms. When higher government spending raises productivity, it increases the scale of production and as a result consumer welfare, which provides a possibility for consumption to increase as well. Linnemann and Schabert (2006) show that even if the impact of public expenditures on production is small, their increase can cause a rise in private consumption if the government share in income is not too large and public finance does not solely rely on distortionary taxation.

Lastly, Linnemann (2006) and Monacelli and Perotti (2008) claim that the positive effect of government spending on consumption may be obtained by choosing a specific form of the utility function. In particular, using a simple real business cycle model, Linnemann (2006) demonstrates that the necessary requirement for the positive consumption response to the government spending shock is a non-separable utility and complementarity between consumption and leisure. At the same time, Monacelli and Perotti (2008) emphasize that the wealth effect on labor supply is important in determining the effect of government spending on consumption. For utility functions where the wealth effect on labor supply is absent, they use a DSGE setting with nominal rigidities to show that consumption increases in response to a government spending shock, while a drop in consumption is observed in a model where the wealth effect on labor supply is large.

In this paper, the focus is on the quantitative comparison of five models - the deep habits model, the model with rule-of-thumb consumers, the model where government spending influences individual preferences directly, the model with productive government expenditures, and finally the baseline model that does not rely on any of these mechanisms. For the baseline model, we adopt a non-separable utility function, which allows for the possibility of either a positive or negative response of consumption to the government spending shock. Because we want to make all models comparable, we use the same type of utility in the other four models as well. While models incorporating these distinct mechanisms have been estimated in separate studies,[2] they normally have variations in model assumptions and data sets, which makes comparisons difficult or impossible. For proper model comparison, we embed the transmission mechanisms into identical frameworks, and estimate them using identical data sets and prior distributions for all common parameters in the models.

We use the Bayesian approach in order to evaluate the relative quantitative performance of these models. While the models have been, in some studies, estimated by matching impulse responses of the government spending shock, we intentionally choose to rely on the full information Bayesian estimation approach. This choice is motivated in part by the controversy that still exists about the response of consumption to a spending shock. It has been established that structural VAR (SVAR) models that utilize timing restrictions through Cholesky decomposition for shock identification generally predict positive co-movement of private and public consumption (see Blanchard and Perotti (2002) or Fatas and Mihov (2001)), while the opposite result is obtained in models where the shock is identified using the narrative approach (see Ramey (2011)). Estimation by matching model impulse responses to the data is subject to conditioning on the shock identification procedure. While it may still be useful as an exercise to verify the ability of a model to produce the positive co-movement, such an estimation strategy does not contribute to the dispute regarding the qualitative response of consumption. One goal of this paper is to find out whether, when taken to match the data unconditionally, the proposed transmission mechanisms predict positive or negative response of private consumption to the government spending shock.

The main result is the finding that no model outperforms the baseline model in terms

[2]For instance rule-of-thumb consumers have been explored by a series of papers featuring an estimated medium scale DSGE model, e.g Forni, Monteforte, and Sessa (2009), Coenen and Straub (2005) and Cogan, Cwik, Taylor, and Wieland (2010) among many others. Zubairy (2014b) explores deep habit formation, Traum and Yang (2015) incorporate productive government capital and Bouakez and Rebei (2007) consider government spending in the utility function in an estimated DSGE model.

of data fit. Therefore, the inclusion of any of the different transmission mechanisms for government spending considered may not be necessary to account for the effect of government spending on consumption. Moreover, differently from the predictions of the SVAR literature, we find that all five of the estimated models consistently generate a negative response of consumption to the government spending shock. In addition, the resulting output multipliers on impact are at or slightly below 1, which is primarily driven by the negative response of consumption to a government spending shock.

In addition we also apply the DSGE-VAR methodology for each model, (see Del Negro and Schorfheide (2004)), in order to examine if the results from the more constrained DSGE models are different from the more flexible DSGE-VAR. The ranking of the various propagation mechanisms is preserved. However, the impact response of consumption to a government spending shock is positive in the DSGE-VAR framework under all the various models, influenced by the Blanchard and Perotti (2002) shock identification in the VAR. This points towards issues of potential misspecification in the DSGE models commonly employed in the study of fiscal policy. For a researcher with strong prior beliefs in the validity of Blanchard and Perotti (2002) identification this suggests that the DSGE models misrepresent the dynamics of consumption in response to a government spending shock, and none of the transmission mechanisms considered work towards resolving this problem. On the other hand, since government spending is assumed to be pre-determined in the DSGE models to be consistent with this identification scheme, the source of mis-specification might be the assumed government spending process. Both Fève, Matheron, and Sahuc (2013) and Kormilitsina (2017) show evidence for government spending responding to output contemporaneously and that excluding it can bias the consumption response to a spending shock.

There are few other papers in the literature that also compare various structurally estimated models to study government spending shocks. For instance, Fève, Matheron, and Sahuc (2013) evaluate responses to government spending shocks and model fit from models with rule of thumb agents and models with government spending in the utility function. The main focus of that paper, however, is examining how the spending multiplier can be biased when endogeneity of government expenditures is disregarded. Another example is Cantore, Levine, and Melina (2014), which compares models with deep and superficial habits in consumption, and emphasizes the importance of assuming persistence in the habit stock. In contrast, our paper undertakes a more comprehensive analysis by considering five different transmission mechanisms for government spending shock in a unified framework.

The paper proceeds as follows: We describe the general framework and model specifics in Section 2. Section 3 offers the strategy for estimation and model analysis. Section 4 discusses estimation results and finally, Section 5 concludes.

2 Models of Government Spending

In this section, we describe the models with distinct propagation mechanisms for government spending shocks used in the quantitative analysis. All these models have some features in common. In particular, each model introduces three types of agents: households, firms and policy authorities. Although exact specification may be different across the models, we assume household's preferences are influenced by consumption habits. We make the same assumptions regarding investment adjustment costs, and endogenous depreciation, which is tied to the degree of capital utilization. The role of money is motivated by nominal price and wage rigidities, while monetary policy is described by a standard Taylor-type rule. In addition to the government spending shock, there are seven other sources of uncertainty. They are the neutral and investment specific technology shocks, preference shock, wage and price markup shocks, tax shock and monetary policy shock. We model the economy as evolving along the balanced growth path, where the long-run trend for consumption, output, wages is different from the long-run trend in capital and investment.

The specific models of government spending extend this set up in the following way: the first model incorporates deep habit formation over consumption of private and public goods. The second model introduces a share of the households being rule-of-thumb consumers. The other two models assume that government spending enhances household utility function or the production technology, respectively. Finally, the baseline model does not have any of these specific features.[3]

2.1 Main Framework

The economy is populated by a continuum of infinitely-lived households. Each household participates in the following activities. It consumes, supplies differentiated labor services to the labor packer, accumulates capital by means of investing, rents capital services to firms, pays taxes and receives dividends from ownership in firms.

2.1.1 Households

Each household derives utility from a consumption measure X_t, the exact definition of which differs across the three models, and homogenous labor h_t. The life-time expected

[3] However, the positive response of consumption is possible because of the wealth effect on labor supply, associated with household utility function that is non-separable between consumption and leisure.

utility of households is defined as

$$E_0 \sum_{t=0}^{\infty} \beta^t d_t U(X_t, h_t),$$

where E_0 denotes expectations based on period zero information set, β is the discount factor, and d_t is the preference shock, evolving according to an $AR(1)$ process:

$$log\left(\frac{d_{t+1}}{d}\right) = \rho_d log\left(\frac{d_t}{d}\right) + \epsilon^d_{t+1}, \tag{1}$$

where $0 < \rho_d < 1$, and $\epsilon^d_t \sim N(0, \sigma_d^2)$, with $\sigma_d > 0$, is an independent and identically distributed (i.i.d.) preference shock. The intratemporal utility function follows King, Plosser, and Rebelo (1988) in that it is nonseparable in leisure and consumption and consistent with long-run balanced growth:

$$U(X_t, h_t) \equiv \frac{X_t^{1-\sigma}}{1-\sigma}(1-h_t)^{\zeta}, \tag{2}$$

where the inverse of $\sigma > 0$ is the inverse of the intertemporal elasticity of substitution in consumption, and ζ is the elasticity of the demand for leisure.[4]

Homogenous labor h_t is a Dixit-Stiglitz aggregate of differentiated labor services h^j_t, for $j \in [0, 1]$:

$$h_t = \left(\int_0^1 (h^j_t)^{1-\frac{1}{\eta^w_t}} dj\right)^{\frac{1}{1-\frac{1}{\eta^w_t}}}.$$

Here, η^w_t is the elasticity of substitution across different types of labor, and the upper script j helps to distinguish between different types of labor. As in Smets and Wouters (2007), the wage markup is modeled as an ARMA (1,1) process,

$$log\left(\frac{\eta^w_{t+1}}{\eta^w}\right) = \rho_w log\left(\frac{\eta^w_t}{\eta^w}\right) + \epsilon^w_{t+1} - \mu_w \epsilon^w_t, \tag{3}$$

where $\eta_w > 1$, $0 < \rho_w < 1$, and $\epsilon^w_t \sim N(0, \sigma_w^2)$, with $\sigma_w > 0$ is an i.i.d. wage markup shock.

In each differentiated labor market j, wages are set by monopolistically competitive unions. The unions determine the nominal wage rate W^j_t to maximize the welfare of the representative consumer, and ensure that the demand for labor type j is fully satisfied. As is

[4]This utility function assumes the existence of wealth effect from the government spending shock on labor supply, as opposed to the GHH type preferences, where this wealth effect is absent. Monacelli and Perotti (2008) show that absent the wealth effect, a standard new-Keynesian model with price rigidities will produce a positive response of consumption to a government spending shock. We choose to avoid the possibility of automatically generating the positive response without eliminating this possibility, allowing the estimation procedure to determine the size of the wealth effect.

standard in the literature, we assume that differentiated labor types are supplied to the labor packer to be aggregated using the Dixit-Stiglitz technology, and the resulting homogenous labor is then supplied to final good producers at an aggregate wage rate W_t. Changes in the wage rate are costly, and modeled as a quadratic adjustment cost

$$\Psi\left(\frac{W_t^j}{W_{t-1}^j}\right) = \frac{\alpha_w}{2}\left(\frac{W_t^j}{W_{t-1}^j} - \mu_{z^*}\pi\right)^2,$$

per (real) dollar of the household's wage bill. In this formula, $\alpha_w > 0$ is the wage adjustment cost parameter, π is the inflation rate along the balanced growth path, and μ_{z^*} is the rate of growth of the economy (output, consumption, and wages) along the balances growth path. The wage adjustment cost is paid by the households in the form of a union membership fee.

The households own physical capital, K_t. Capital is accumulated through the process of investing, and the total stock of capital depreciates at a variable rate depending on how intensively it is used. Moreover, investment adjustment is costly, with the capital loss of $S(\cdot)$ per unit of investment. The dynamics of capital is therefore:

$$K_{t+1} = (1 - \delta(u_t))K_t + I_t\left(1 - \mathcal{S}\left(\frac{I_t}{I_{t-1}}\right)\right), \tag{4}$$

where u_t determines the intensity of capital utilization as a fraction of capital being used in production, and $\delta(u_t)$ is the depreciation function, parameterized as follows:

$$\delta(u_t) = \delta_0 + \delta_1(u_t - u) + \frac{\delta_2}{2}(u_t - u)^2, \tag{5}$$

where $\delta_0, \delta_1, \delta_2 \geq 0$, and u is the steady state rate of capital utilization. In Equation (4), the cost of investment $S(\cdot)$ is the quadratic function:

$$\mathcal{S}\left(\frac{I_t}{I_{t-1}}\right) = \frac{\kappa}{2}\left(\frac{I_t}{I_{t-1}} - \mu_I\right)^2,$$

where $\kappa > 0$, and μ_I is the steady-state growth rate of capital and investment.

Following Fisher (2003), investment goods I_t are obtained from consumption using a stochastic linear technology, according to which at each date t, one unit of consumption can produce Υ_t units of investment. We call Υ_t the investment specific technology. Denoting $\mu_{\Upsilon,t} \equiv \Upsilon_t/\Upsilon_{t-1}$, the gross growth rate of Υ_t, the dynamics for $\mu_{\Upsilon,t}$ is

$$log\left(\frac{\mu_{\Upsilon,t+1}}{\mu_\Upsilon}\right) = \rho_\Upsilon log\left(\frac{\mu_{\Upsilon,t}}{\mu_\Upsilon}\right) + \epsilon_{t+1}^\Upsilon, \tag{6}$$

where $\epsilon_t^{\Upsilon} \sim N(0, \sigma_{\Upsilon}^2)$ is an i.i.d. shock, with $\sigma_{\Upsilon} > 0$, and μ_{Υ} is the growth rate of the investment specific technology along the balanced growth path.

Capital services $u_t K_t$ are rented out to firms at a real rental rate R_t^k. Households own shares in firms, and receive dividends with the real value Φ_t. They pay a distortionary income tax, at the rate τ_t, and receive lump-sum transfers in the amount Tr_t in terms of consumption. Complete set of one-period state-contingent assets, as well as the risk-free government bonds are traded in financial markets. If households have access to financial markets,[5] then the budget constraint can be written in real terms as[6]

$$\begin{aligned} E_t r_{t,t+1} L_{t+1} \quad + \quad & C_t + \Upsilon_t^{-1} I_t + \frac{B_{t+1}}{R_t} = \\ & \frac{L_t}{\pi_t} + (1-\tau_t) R_t^k u_t K_t + \int \left(1 - \tau_t - \Psi\left(\frac{W_t^j}{W_{t-1}^j}\right)\right) W_t^j h_t^j dj + \Phi_t + \frac{B_t}{\pi_t} + Tr_t, \end{aligned}$$

where L_t is the payoff in period t of state-contingent securities traded in period $t-1$, $r_{t,t+1}$ is the price of a state contingent security traded at date t for a claim on consumption delivered in period $t+1$, C_t is real consumption, τ_t is the income tax rate, and B_t is the real value of non-state contingent government bonds in possession of households. The new bonds are purchased at a price $1/R_t$.

2.1.2 Firms

A continuum of monopolistically competitive firms of measure 1 produce differentiated intermediate goods. For production, each firm uses capital and labor services, $u_t K_t$ and h_t according to the following technology

$$F(u_t K_t, Z_t h_t) \leq q_t (u_t K_t)^{\theta} (Z_t h_t)^{1-\theta} - Z_t^* \vartheta, \tag{7}$$

where $0 < \theta < 1$, variable q_t is model specific, introduced in Section 2.2, $Z_t^* \vartheta$ represents the fixed costs of operating a firm in each period,[7] Z_t is the stochastic labor-augmenting productivity process, growing at a rate of $\mu_{z,t}$, $\mu_{z,t} \equiv Z_t / Z_{t-1}$, which evolves according to and $AR(1)$ process:

$$log\left(\frac{\mu_{z,t+1}}{\mu_z}\right) = \rho_z log\left(\frac{\mu_{z,t}}{\mu_z}\right) + \epsilon_{t+1}^z. \tag{8}$$

[5]This is the case in all models except for the model with rule-of-thumb consumers.

[6]To simplify notation, we omit the household specific superscript j when it is possible.

[7]Z_t^* is the stochastic trend for the economy, which is a combination of the investment specific and labor-augmenting technologies.

Here, μ_z is the growth rate along the balanced growth path, $0 < \rho_z < 1$, and $\epsilon_t^z \sim N(0, \sigma_z^2)$, with $\sigma_z > 0$ is an i.i.d. neutral technology shock.

Each firm $i \in [0, 1]$ maximizes the present discounted value of dividend payments, given by

$$E_t \sum_{s=0}^{\infty} r_{t,t+s} P_{t+s}^i \Phi_{t+s}^i,$$

where $r_{t,t+s} \equiv \prod_{k=1}^{s} r_{t+k-1,t+k}$, for $s \geq 1$, with $r_{t,t} \equiv 1$, and period t dividend payments in real terms are

$$\Phi_t^i = \frac{P_t^i}{P_t} a_t^i - R_t^k u_t^i K_t - W_t h_t^i - \Omega\left(\frac{P_t^i}{P_{t-1}^i}\right),$$

where a_t^i is the demand for the firm i's output, $\Omega(\cdot)$ is the cost of price changes, following Rotemberg (1982). We assume that this cost is quadratic and proportional to the stochastic trend Z_t^*:

$$\Omega\left(\frac{P_t^i}{P_{t-1}^i}\right) = \frac{\alpha_p Z_t^*}{2} \left(\frac{P_t^i}{P_{t-1}^i} - \pi\right)^2,$$

with $\alpha_p > 0$, denoting the degree of price stickiness. Monopolistically competitive firms must satisfy their demands at the posted price.

In all models except the one with deep habits, the good intended for final consumption is the aggregate of differentiated goods produced by monopolistically competitive firms using a Dixit-Stiglitz technology:

$$\left(\int_0^1 (Y_t^i)^{1-\frac{1}{\eta_{p,t}}} di\right)^{\frac{1}{1-\frac{1}{\eta_{p,t}}}}.$$

where $\eta_{p,t}$ is the elasticity of substitution between individual good varieties. which is assumed to follow an ARMA (1,1) process,

$$log\left(\frac{\eta_{t+1}^p}{\eta^p}\right) = \rho_p log\left(\frac{\eta_t^p}{\eta^p}\right) + \epsilon_{t+1}^p - \mu_p \epsilon_t^p, \tag{9}$$

where $\eta_t^p > 1$, $0 < \rho_p < 1$, and $\epsilon_t^p \sim N(0, \sigma_p^2)$, with $\sigma_p > 0$, is an i.i.d. price markup shock.

2.1.3 Fiscal and monetary policy

The fiscal authority levies taxes, provides lump-sum transfers and develops public projects with real cost of G_t. We assume that each period, the government satisfies a balanced budget. To ensure the model has a well-defined balanced growth path, we assume that government expenditures G_t evolve along the same stochastic trend as output and consumption. With this purpose, we assume the ratio $\varsigma_t^g = G_t/Y_{t-1}$ is an AR(1) process:[8]

$$log\left(\frac{\varsigma_{t+1}^g}{\varsigma^g}\right) = \rho_g log\left(\frac{\varsigma_t^g}{\varsigma^g}\right) + \epsilon_{t+1}^g, \tag{10}$$

where $0 < \rho^g < 1$, and $\epsilon_t^g \sim N(0, \sigma_g^2)$, with $\sigma_g > 0$, is an i.i.d. government spending shock.

Households face distortionary taxes on their income to finance government spending, and the income tax rate τ_t evolves according to the following process:

$$log\left(\frac{\tau_t}{\tau}\right) = \alpha_\tau log\left(\frac{\tau_{t-1}}{\tau}\right) + \alpha_{\tau,y}\log(\frac{Y_{t-1}}{\tilde{Y}_{t-1}}) + \epsilon_t^\tau, \tag{11}$$

where $0 < \alpha_\tau < 1$, and $\epsilon_t^\tau \sim N(0, \sigma_\tau^2)$, with $\sigma_\tau > 0$, is an i.i.d. tax shock. $\tilde{Y}_t = Z_t^* y$, where y is the steady state level of detrended output. Parameter $\alpha_{\tau,y}$ measures the response of the tax rate to economic conditions measured by the output gap, which captures automatic stabilizer effects.[9]

We assume that monetary policy is described by a generalized Taylor type rule with the interest rate smoothing and response to inflation and output growth, as follows:

$$log\left(\frac{R_t}{R}\right) = \alpha_R log\left(\frac{R_{t-1}}{R}\right) + \alpha_\pi log\left(\frac{\pi_t}{\pi}\right) + \alpha_Y log\left(\frac{Y_t}{Y_{t-1}\mu_{z^*}}\right) + \epsilon_t^r, \tag{12}$$

where Y_t is aggregate real output, α_R, α_π, α_Y are Taylor rule parameters, and $\epsilon_t^r \sim N(0, \sigma_r^2)$ is an i.i.d monetary policy shock, with $\sigma_r > 0$.

2.2 Model Specific Features

In this section, we briefly describe the five models we consider, with an emphasis on their specific features. More details on the models, including the first order and market clearing conditions, are given in the Technical Appendix accompanying the paper, available online.

[8] Such modeling assumption is motivated by the fact that planned government expenditures are decided upon prior to the quarter of implementation, and therefore current public expenditures are predetermined with respect to current output. This is similar to the identification assumption in Blanchard and Perotti (2002). To implement this idea, we define ς_t^g as the ratio of G_t and the previous, rather than current period output, Y_{t-1}. We find that this approach to modeling government spending shock improves the marginal likelihood of all models. Alternatively, within-period timing restrictions could be imposed, as in Kormilitsina (2013).

[9] The consequences of different methods of financing, including distortionary taxes are emphasized by Leeper, Plante, and Traum (2010).

2.2.1 Model with Deep Habits

We adopt the "fully-fledged" version of the deep habits model from Ravn, Schmitt-Grohé, and Uribe (2006), according to which both private and public consumption is subject to deep habits. With deep habits, effective consumption X_t in Equation (2) is defined as

$$X_t = \left[\int_0^1 (C_{i,t} - b^c S^c_{i,t-1})^{1-\frac{1}{\eta_{p,t}}} di \right]^{1/(1-\frac{1}{\eta_{p,t}})},$$

where index i refers to a variety of differentiated goods produced by monopolistically competitive firms, b^c is the habit formation parameter for private consumption.[10]

$S^c_{i,t}$ is the good-specific stock of habit, which evolves over time according to the law of motion,

$$S^c_{i,t} = \rho^c S^c_{i,t-1} + (1-\rho^c) C_{i,t},$$

with $0 \leq \rho^c \leq 1$.

Similar to Ravn, Schmitt-Grohé, and Uribe (2006), the government allocates spending over intermediate goods $G_{i,t}$ so as to maximize the quantity of a composite good X^g_t produced with intermediate goods according to the relationship

$$X^g_t = \left[\int_0^1 (G_{i,t} - b^g S^g_{i,t-1})^{1-\frac{1}{\eta_{p,t}}} \right]^{1/(1-\frac{1}{\eta_{p,t}})},$$

where b^g is the habits parameter for public goods, and the stock of habits $S^g_{i,t}$ is determined as follows

$$S^g_{i,t} = \rho^{gg} S^g_{i,t-1} + (1-\rho^{gg}) G_{i,t},$$

where $0 \leq \rho^{gg} \leq 1$.

Parameter q_t of the production function in Equation (7) is set to 1.

2.2.2 Model with Rule-of-Thumb Consumers

As in Galí, Lopez-Salido, and Vallés (2007), we assume that only a fraction $(1-\lambda)$ of all households have access to capital markets where they can trade state-contingent bonds and accumulate capital to rent out to firms. These are known as optimizing households. Other households, the so-called rule-of-thumb consumers, do not participate in financial markets, therefore they cannot borrow or save. These households are restricted to consume their entire disposable labor income.

[10]Note that deep habits imply habit formation at the level of the intermediate good $C_{i,t}$, and not the aggregate good C_t that enters the utility function.

Utility of optimizing households is determined by Equation (2), where X_t is the habit adjusted consumption,

$$X_t^o = C_t^o - h^c C_{t-1}^o,$$

in which h^c is the consumption habits parameter and C_t^o denotes homogenous consumption of optimizing households at date t.

The labor unions ensure both optimizing and rule-of-thumb households work sufficient hours in order to meet the market demand at the wage rate set by a monopolistic competitive union. We follow Galí, Lopez-Salido, and Vallés (2007) in assuming that the rule-of-thumb households work exactly the same hours (h_t^r) as the optimizing consumers (h_t^o):

$$h_t^r = h_t^o \equiv h_t.$$

As mentioned earlier, the differentiated labor unions maximize the welfare of their representative member, which is given by the weighted average of the present values of expected life time utilities:

$$E_0 \beta^t d_t [\lambda U(X_t^c, h_t) + (1-\lambda) U(X_t^o, h_t)],$$

where X_t^c is defined analogously to X_t^o. In a symmetric equilibrium, with the rule-of-thumb consumers providing differentiated labor services, the wage rates for both types of households coincide, thus $W_t^r = W_t^o = W_t$ at any period t.[11]

Consumption of the rule-of-thumb households is determined after income taxes and wage cost adjustment as follows:

$$C_t^r = \left(1 - \tau_t - \Psi\left(\frac{W_t}{W_{t-1}}\right)\right) W_t h_t + Tr_t^r.$$

In this formula, $Tr_t^r = Z_t^* \tau^r$ is the lump-sum transfer that the rule-of-thumb households receive from government. We assume the detrended transfer, τ^r, is constant.[12] To close the model properly, we assume that the lump sum tax on optimizing households varies in order to ensure that the government budget is balanced in each period.

Finally, parameter $q_t \equiv 1$ in Equation (7).

[11]Note, that each member of the union, optimizing or rule-of-thumb consumer, is subject to the wage adjustment cost which can be thought of as an equal share of a union membership fee.

[12]The steady state value for τ^r is pinned down by the equality of steady state consumption for optimizing and rule-of-thumb households.

2.2.3 Model with Government Spending in the Utility Function

We follow Bouakez and Rebei (2007) and define X_t in the intratemporal utility in Equation (2) as habit adjusted effective consumption,

$$X_t = \tilde{C}_t - h^c \tilde{C}_{t-1},$$

where $h^c > 0$ is the habit formation parameter, and the effective consumption, $\tilde{C}_t$, is a combination of private and public consumption, C_t and G_t:

$$\tilde{C}_t = \left[\phi C_t^{\frac{\nu-1}{\nu}} + (1-\phi) G_t^{\frac{\nu-1}{\nu}} \right]^{\frac{\nu}{\nu-1}},$$

where $0 < \phi < 1$ is the weight on private consumption in the effective consumption, and $\nu \geq 0$ is the elasticity of substitution between private and public spending. When $\nu \to 0$, private and public consumption are nearly perfect complements and they become substitutes as $\nu \to \infty$.

Parameter $q_t \equiv 1$ in the production function defined in Equation (7).

2.2.4 Model with Productive Government Spending and the Baseline Model

One difference between these two models is that in the baseline model, $q_t \equiv 1$ in Equation (7), while in the model with productive government spending, we acknowledge that government actions may directly affect the production process. Similar to Baxter and King (1993), public spending enhances the production technology in Equation (7) through q_t.

For the model with productive government spending only, we distinguish between public consumption and capital. We assume that public capital improves private sector production technology by affecting q_t in the following way:

$$q_t = \left(\frac{K_t^g}{Z_t^*} \right)^{\alpha_G},$$

where $\alpha_G > 0$, and K_t^g represents public capital. We assume that government investment expenditures I_t^g contribute to public capital accumulation (K_t^g) according to the dynamic equation

$$K_{t+1}^g = (1 - \delta(u_t)) K_t^g + I_t^g.$$

Total government spending G_t consists of public consumption and investment, which are constant shares of G_t:

$$G_t^c = s^{gc} G_t,$$

$$I_t^g = (1 - s^{gc})G_t,$$

where s^{gc} denotes the share of public consumption.

Utility in both models features standard superficial habit in consumption, therefore X_t in Formula (2) is defined as

$$X_t = C_t - h^c C_{t-1},$$

where h^c is the consumption habit parameter, and C_t is consumption of final goods.

2.3 Propagation Mechanisms of the Government Spending Shock

According to a standard real business cycle (RBC) model, the government spending shock reduces resources of the economy generating a negative wealth effect. As a result, consumption falls, while output and labor increase. While it is widely accepted that a rise in government spending stimulates production and employment, its negative effect on consumption is puzzling in light of some of the empirical evidence. Bilbiie (2009) demonstrates that in a simple RBC framework, there is no possibility for consumption to rise in response to rising government spending, unless the labor supply function is negatively sloped. This can be demonstrated graphically using Figure 1. The figure shows the equilibrium in the market for labor services. In the figure, the real wage rate (w) is plotted along the vertical, and labor hours (h) along the horizontal axis. The thick solid line in the figure represents the supply of labor before the shock, while the thick starred line is the labor demand of firms. The supply of labor is determined by,

$$w = \frac{U_2(c, 1-h)}{U_1(c, 1-h)}. \tag{13}$$

The labor demand is given by,

$$w = mcF_h(uk, h), \tag{14}$$

where mc is the real marginal cost of firms, which is the inverse of firm's price markup over the marginal costs. In the standard RBC framework, mc, and u equal 1 at all times and in all states of the economy. The marginal rate of substitution between consumption and leisure, $\frac{U_2(c,1-h)}{U_1(c,1-h)}$, is usually increasing both in consumption and labor factor. This property ensures that the labor supply is positively sloped, and a drop in consumption shifts the labor supply to the right, while an increase in consumption shifts the labor supply to the left.

In a standard RBC model, a rise in government spending is associated with the negative wealth effect, and thus it causes a drop in consumption and a rise in labor supply. Note from Equation (13) that labor supply increases because when consumption drops, the same real wage rate is associated with a larger supply of labor. Therefore, the equilibrium moves instantaneously from point 0 to 1 in Figure 1. If an equilibrium increase in consumption were a possibility in this model, this would cause labor supply to decrease according to Equation (13), shifting the labor supply curve to the left, with the new equilibrium at point 2. However, this scenario is not supported by equilibrium in the most standard version of the model, because this move would reduce equilibrium labor; therefore output would shrink leaving no possibility for consumption to expand. As a result, the necessary condition for consumption to rise is that the new equilibrium supports larger employment, allowing output to expand. Such a possibility may arise in a model where labor demand increases endogenously due to rising government spending. This scenario is shown by point 3 in Figure 1. At this point, labor is larger than that at point 0 as well as consumption, which means that equilibrium with higher consumption level is supported at this point.

The introduction of imperfectly competitive goods market and price stickiness allow the labor demand to increase as well. With price stickiness, an increase in output demand due to the rising government spending is associated with larger marginal costs and increased labor demand, since firms can not adjust prices easily. Because the marginal cost, mc, is the inverse of the firm's markup, mc and output move pro-cyclically.

It has been shown however, that price stickiness, cannot generate sufficiently large shifts in the labor demand curve to guarantee a rise in consumption in response to the government spending shock (see Linnemann and Schabert (2003)). Therefore, additional assumptions are needed to overturn the negative wealth effect on consumption. In the model with rule of thumb consumers, this is done by introducing a fraction of non-Ricardian consumers who consume their entire disposable income in every period, following the so called "rule of thumb", or because they have no access to financial markets. Because optimizing households still experience a drop in wealth due to a rise in public spending, the rise in total consumption can only be achieved if rule-of-thumb households increase consumption substantially, which can only happen if the wage income of these non-optimizing households increases. The wage income rises if the wage rate or hours worked increase. Because optimizing agents demand to work more, the wage rate tends to drop. Galí, Lopez-Salido, and Vallés (2007) rely on an important assumption that labor markets are non-competitive in a specific way so that both types of households always work the same hours. This assumption allows the labor of rule-of-thumb households to increase when government spending rise, therefore making rising income, and consequently consumption, a possibility. Certainly, aggregate consumption in this model will only increase if the share of rule-of-thumb consumers, λ, is large enough to compensate for the drop in consumption of optimizing households.[13]

[13] While the model considered in Galí, Lopez-Salido, and Vallés (2007) is a standard new-Keynesian frame-

The mechanism of the deep habits follows the same route as that of nominal price rigidity, because it works by generating endogenous countercyclical markups of firms.[14] The reason is that the combination of deep habits and imperfect competition results in time-varying elasticity of demand. To see this, note that in a simplified deep habits model, the demand for consumption good i is given by[15]

$$C_{i,t} = \left(\frac{P_{it}}{P_t}\right)^{-\eta_{p,t}} (C_t - b^c C_{t-1}) + b^c C_{i,t-1}.$$

For this demand function, the price elasticity is $\eta_{p,t}\left(\frac{P_{it}}{P_t}\right)^{-\eta_p} \frac{(C_t - b^c C_{t-1})}{C_{i,t}}$,[16] which is proportional to the habit adjusted aggregate consumption level, $(C_t - b^c C_{t-1})$. Therefore, when aggregate consumption rises, the price elasticity of demand increases, and everything else equal, producers have incentives to reduce markups. By doing so, firms gain a larger share of the market to form the stock of habits and increase future profits. The resulting drop in markups raises mc, and therefore increases the labor demand curve in the same way as in the mechanism with sticky prices. However, Ravn, Schmitt-Grohé, and Uribe (2007) show that deep habits can generate much larger movements in the markups and consequently labor demand, than price stickiness, providing a better foundation to obtain the positive response of consumption to the government spending shock. These authors show that the deep habits mechanism helps explain a rise in consumption even in the absence of price stickiness.[17]

The similar outward shift in the demand for labor leading to the new equilibrium in point 3 occurs in the model with productive government spending. In this case, however, the labor demand curve shifts out due to a rise in labor productivity, F_h in Equation (14), rather than the marginal cost. If the effect of government spending on labor productivity is large enough, then the rise in consumption may be an equilibrium outcome.

According to Linnemann and Schabert (2004) and Bouakez and Rebei (2007) if private

work, Furlanetto (2011) and Colciago (2011) show that the introduction of wage rigidities does not change this result qualitatively. Nominal wage rigidities mitigate the fall in the wage rate, reducing the negative wealth effect on optimizing households, and may also increase the disposable income of rule-of-thumb households. Thus, strong nominal wage stickiness may guarantee the positive correlation between public and private consumption for the rule-of-thumb households.

[14]This counter-cyclicality of markups has been documented in many empirical studies, such as Bils (1987), Rotemberg and Woodford (1991) and Rotemberg and Woodford (1999).

[15]This is the simplified demand function under the assumption of $\rho^c = 0$. Note also that in the absence of deep habits, the demand function would be, $C_{i,t} = (P_{it}/P_t)^{-\eta_p} C_t$, implying time-invariant price elasticity of η_p.

[16]A similar demand function holds for the intermediate government spending good, G_{it} and there is similar intuition behind pro-cyclicality of price elasticity in response to increased demand from government spending.

[17]In the presence of both price stickiness and deep habits, there can be interesting interactions, as shown in Jacob (2013), where an increasing level of price stickiness mitigates the crowding-in effect on consumption in response to a spending shock.

and public consumption are complements in the sense that an increase in government spending raises marginal utility of consumption, then a rise in government spending increases labor supply as shown in Figure 2. With this move, an equilibrium with rising consumption response becomes a possibility, because it does not have to be associated with a reduction in labor supply, as shown by the new equilibrium at point 5 in the figure.

The form and calibration of the utility function by itself plays an important role in the resulting effect of government spending shock on consumption. Linnemann (2006) claims that in the RBC setting, the necessary condition for a rise in consumption is that consumption and leisure must be substitute goods in the sense that $U_{12} < 0$.[18] Monacelli and Perotti (2008) emphasize the importance of the wealth effect on labor supply in determining the response of consumption to the government spending shock. The idea there is that the smaller is the shift of the labor supply curve as a result of the shock, the more likely the new equilibrium will move north-east of point 0 in Figure 1, raising both wages and hours.[19] In the example they use, consumption rises in the economy with nominal price stickiness where preferences feature no wealth effect on labor supply, and fall if preferences are such that the wealth effect on labor supply is significant.

The models we estimate have features commonly used in estimated DSGE models, such as nominal wage rigidities, habit formation, investment adjustment cost, and endogenous capital utilization. While being helpful in achieving better fit with data, these features complicate intuition behind the propagation mechanism of the government spending shock. Nevertheless, we should expect that rigid wages mitigate fluctuations in income resulting from the government spending shock, therefore reducing the wealth effect on labor supply, and increasing the possibility to observe positive consumption response to the shock. Introducing superficial habits have consequences for the labor supply curve since habits will affect the wealth effect on labor supply and the resulting consumption behavior. Monacelli and Perotti (2008) demonstrate that adding habits to the simple RBC model without price stickiness helps in obtaining a positive response of consumption to a government spending shock. Endogenous capital utilization makes it possible for the labor demand to respond endogenously to rising government spending even in the standard RBC setting. Although response of capital utilization to the shock is endogenously determined, it is expected to increase when public spending rises, affecting the demand for labor in a way similar to price stickiness. All in all, the presence of these features to some extent will have a quantitative influence on the consumption effect of government spending shocks.

[18] With nominal rigidities, however, this does not have to be the case.

[19] Notice that larger wages are desirable for the positive effect of government spending on consumption, because of the consumption leisure substitution effect they create - the larger is the real wage rate the less expensive is consumption, making it more attractive for households.

3 Estimation and Inference

3.1 Estimation Strategy

The models we study can be cast in a linear state space form, with athe likelihood derived via a Kalman filter, which when coupled with priors on model parameters delivers posterior distribution for the parameter vector θ conditional upon the model. In doing so we keep the data employed in the observable equation constant across the models. The data y_t is the 8×1 vector of observable variables defined as follows

$$y_t = \{\, log(G_t/Y_{t-1}),\, log(Tax_t/Y_t),\, \Delta(log(C_t)),\, \Delta(log(I_t)),\, log(H_t),\, \Delta(log(W_t)),\, 4\Delta(log(P_{Y,t})),\, R_t \,\},$$

where G_t/Y_{t-1} and Tax_t/Y_t give the government spending to GDP ratio, and the tax revenues to GDP ratio, respectively. C_t, I_t, and W_t are real per capita real consumption and investment, and real wages, which appear as growth rates to be consistent with their model implied nonstationarity property. H_t are per capita hours worked, $P_{Y,t}$ is GDP deflator, and therefore $4\Delta(log(P_{Y,t}))$ measures annualized inflation rate based on the GDP deflator. R_t is the nominal interest rate, measured by the effective (annualized) Federal funds rate.[20] All the data in vector y_t are expressed as deviations from their means, and appear in quarterly frequency, spanning 1954:3 to 2010:4.

The vector of estimated model parameters is defined as

$$\theta = \{\, \theta^1_{A_i},\, \theta^2,\, \theta^3 \,\},$$

where $\theta^1_{A_i}$ is the vector of model A_i specific parameters, for models $i = \{DH, ROT, UTIL, PROD\}$ θ^2 is the vector of parameters common across the models, and θ^3 is the vector of parameters calibrating the shock processes. These three groups of parameters consist of the following elements:

$$\theta^1_{DH} = \{\, b^c,\, \rho^c,\, b^g,\, \rho^{gg} \,\}, \qquad \theta^1_{ROT} = \{\, h^c,\, \lambda \,\}, \qquad \theta^1_{UTIL} = \{\, h^c,\, \nu,\, \phi \,\}, \qquad \theta^1_{PROD} = \{\, h^c,\, \alpha_G \,\},$$

[20] G_t is given by total government consumption expenditures and gross investment, which includes federal, state and local spending. This choice is motivated by the fact that this measure of government spending is typically used in SVAR literature when identifying spending shocks. Tax_t are the tax revenues computed as the sum of personal current taxes, taxes on corporate income and contribution for government social insurance and Y_t denotes GDP. C_t is private consumption of nondurable goods and services, I_t is calculated as the sum of durable consumption and private investment. H_t denotes non-farm business sector hours of all persons, and W_t denotes non-farm business sector compensation per hour. The real per capita variables are obtained dividing by labor force and the GDP deflator, $P_{Y,t}$. The data for output and its components, and tax revenues are obtained from BEA, the data for the labor force and hours and wages are from the BLS and the Federal funds rate data is from St. Louis FRED.

$$\theta^2 = \{\, \alpha_p,\, \alpha_w,\, \kappa,\, \delta_2/\delta_1,\, \sigma,\, \alpha_R,\, \alpha_\pi,\, \alpha_Y,\, \alpha_\tau,\, \alpha_{\tau,y} \},$$

$$\theta^3 = \{\, \rho_g,\, \rho_z,\, \rho_\Upsilon,\, \rho_d,\, \rho_p,\, \rho_w,\, \sigma_g,\, \sigma_z,\, \sigma_\Upsilon,\, \sigma_d,\, \sigma_r,\, \sigma_p,\, \sigma_w,\, \sigma_\tau,\, \mu_p,\, \mu_w \,\}.$$

Parameters presented in Table 1 are calibrated, either because it is conventional in the literature, or because estimating these parameters is problematic due to identification issues. The parameter governing the steady state share of capital is set at $\theta = 0.3$. Following Altig, Christiano, Eichenbaum, and Linde (2011), the steady state growth rate of output, μ_{z^*}, is calibrated at 1.0047, while the growth rate of the embodied technology is set at 1.0042. The steady state gross rate of inflation is calibrated as $\pi = 1.0086$, to match the average yearly rate of inflation of 3.5 percent. The intertemporal discount factor $\beta = 0.999$. This relatively high value for β ensures the steady state nominal interest rate is below 6 percent, because smaller values for β implies unrealistically large steady state nominal interest rates. The steady state rate of capital utilization is $u = 1$, while the steady state depreciation rate is fixed at a conventional value $\delta_0 = 0.025$. The actual average share of government expenditures in GDP, $G/Y = 0.2$, is used to calibrate the steady state share of government expenditures in the model. Finally, we fix the elasticity of substitution for intermediate goods and labor types, because estimating these parameters is problematic. We set η_p at 6 and η_w at 21, which imply the steady state price and wage markups of 20 and 5 percent correspondingly.

Column 2 of Tables 4 and 5 show the prior distribution of the estimated parameters in the five models. These distributions are chosen from beta, gamma or inverse gamma distributions. All parameters with bounded support have a beta prior, while gamma and inverse gamma distributions are chosen as priors for parameters bounded from below, such as parameters of the nominal rigidities, investment costs, and others. The priors for these parameters are centered at different values, dictated by the common knowledge generated by the empirical literature. The prior distribution for standard deviations of shock processes are uniform distributions.

3.2 Model Comparison

In order to evaluate the relative quantitative performance of the models, we estimate and compare their marginal likelihoods. Suppose $Y_T = \{y_t\}_{t=1}^T$ is the observed history of vector y_t up to period T, and $Y_0 = \emptyset$. The posterior probability of model A_i is determined by the Bayes formula

$$p(A_i|Y_T) = P(A_i)p(Y_T|A_i),$$

where $P(A_i)$ is the prior probability, and $p(Y_T|A_i)$ is the marginal probability of Y_T, or the likelihood function. For any two models, A_i and A_j, the posterior odds ratio is defined as

$$\frac{p(A_i|Y_T)}{p(A_j|Y_T)} = \frac{P(A_i)}{P(A_j)} \left[\frac{p(Y_T|A_i)}{p(Y_T|A_j)} \right], \tag{15}$$

where $\frac{P(A_i)}{P(A_j)}$ is the ratio of prior probabilities of the two models, called the prior odds ratio, and $\left[\frac{p(Y_T|A_i)}{p(y_T|A_j)} \right]$ is the ratio of marginal likelihoods of the two models, or *the Bayes factor*. Denoting $L(i|j)$ as the loss incurred if choosing model A_i when model A_j is true, the expected posterior loss from choosing model A_i is $P(A_j|Y_T)L(i|j)$. Then, one should choose model A_i if the expected posterior loss from choosing it is smaller than that of the alternative model, or $P(A_j|Y_T)L(i|j) < P(A_i|Y_T)L(j|i)$. This expression can be rewritten as follows

$$\frac{p(A_i|Y_T)}{p(A_j|Y_T)} > \frac{L(i|j)}{L(j|i)},$$

the right hand side of which is usually called the Bayes critical value. Model A_i should be preferred to model A_j if the posterior odds ratio exceeds the Bayes critical value. Combining this expression and Equation (15), one can obtain that

$$\frac{p(Y_T|A_i)}{p(Y_T|A_j)} > \frac{L(i|j)}{L(j|i)} \frac{P(A_j)}{P(A_1)}.$$

If the researcher has prior beliefs about the validity of the two models, and is able to evaluate the relative cost of making a mistake regarding what the true model is, then the posterior odds ratio will provide enough information to choose the model that better explains the data Y_T. When there is no strong evidence regarding the prior odds or the Bayes critical value, it is reasonable to set $L(i|j) = L(j|i)$, and $P(A_i) = P(A_j)$. In this case, the model with the larger marginal likelihood should be chosen as the preferred model.

Since we do not want to create a bias in favor of any model, we assume all five models have equal prior probabilities, and the same expected posterior losses. We thus compare the models' marginal likelihoods, and leave it to the readers to adjust the reported results about the best fitted model using their prior beliefs.

To calculate the model's marginal likelihood, we implement the Harmonic mean estimator of Gelfand and Dey (1994), described in detail by Geweke (1999). Gelfand and Dey notice that for any p.d.f. $f(\theta)$ with the support in Θ, the posterior mean of

$$\frac{f(\theta)}{p(\theta|A_i)p(Y_T|\theta, A_i)} \tag{16}$$

coincides with the inverse of the marginal likelihood of the model:

$$E\left[\frac{f(\theta)}{p(\theta|A_i)p(Y_T|\theta, A_i)}|Y_T, A_i\right] = P^{-1}(Y_T|A_i).$$

Suppose the support of $f(\theta)$ is $\hat{\Theta}_M = \{\theta : (\theta - \hat{\theta}_M)'\hat{\Sigma}_M^{-1}(\theta - \hat{\theta}_M) \leq \chi^2_{1-p}(k)\}$, where p is any number on interval $(0, 1)$, $\hat{\theta}_M = \frac{\sum_{m=1}^{M}\theta^{(m)}}{M}$ and $\hat{\Sigma}_M = \frac{\sum_{m=1}^{M}(\theta^{(m)}-\hat{\theta}_M)(\theta^{(m)}-\hat{\theta}_M)'}{M}$, and $\chi^2_{1-p}(k)$ is the p-value of the χ^2 distribution with k degrees of freedom. Geweke (1999) shows that $f(\theta)$ defined on $\hat{\Theta}_M$ as

$$f(\theta) = p^{-1}(2\pi)^{-k/2}|\hat{\Sigma}_M|^{-1/2}exp[-(1/2)(\theta - \hat{\theta}_M)'\hat{\Sigma}_M^{-1}(\theta - \hat{\theta}_M)], \tag{17}$$

will guarantee the boundedness of expression (16), and thus the posterior mean will exist as long as the posterior density $p(\theta|Y_T, A_i)$ is uniformly bounded away from zero on every compact subset of Θ.

To calculate the posterior expectation of the expression in (16), we evaluate the mean value of the elements of the Markov chain used to calculate the parameter estimate. As noted in Geweke (1999), the estimator may sometimes be quite unstable. To confirm the stability of our results, we compute and report the marginal likelihood for different values of p.

Besides reporting the relative model fit, we also compare models in terms of their predictive ability by evaluating predictive likelihoods in the spirit of Geweke and Amisano (2010). The predictive likelihood function is the model's probability density for the observable variables at the relevant horizon before they are observed, evaluated at the actual values after they are observed. More specifically, for each model j a 1-period ahead predictive likelihood is

$$PL_{A_j}(t) = p(y_t|Y_{t-1}, A_j).$$

It is easy to see that one can evaluate $PL_{A_j}(t)$ as the ratio of marginal likelihoods

$$PL_{A_j}(t) = \frac{p(Y_t|A_j)}{p(Y_{t-1}|A_j)}.$$

This requires calculating marginal likelihood using different datasets, which implies additional estimations for the parameter vector. Because repeated estimation of 5 models is computationally expensive, we evaluate the predictive likelihoods for five-year intervals, and report them as average quarterly values.

4 Estimation Results

4.1 Model Comparison

The results of the model comparison exercise are presented in Table 2. The first column indicates the value of p from Formula (17), used to calculate the log marginal likelihood of the models. Column 2 provides the estimate of the log marginal likelihood for the baseline model. Columns 3 - 6 show marginal likelihood less than that of the baseline model; therefore negative numbers indicate poorer fit of a model.

First of all, Table 2 reveals that the resulting model marginal likelihood values are very similar for all values of p. The log marginal likelihood of the baseline model (BL) is the largest, and varies around 5094 depending on the value of p. The models with government spending in the utility function and production demonstrate just a slightly more inferior fit, with the log Bayes factor taking the largest value of 7.2 and 4.4 in favor of the baseline model. Such difference between these models is "strongly" in favor of the baseline model, according to the classification in Jeffreys (1961). The log marginal likelihood of the rule-of-thumb consumer model and deep habits model are even larger and translate into a Bayes factors of e^{22} and e^{34} in favor of the baseline model, respectively, each of which is significantly greater than 1000. Therefore, Bayesian comparison is decisive in favor of the baseline model. Thus, Table 2 identifies the baseline model as the one with the best performance at describing the data.[21]

To shed more light on the models' relative predictive ability, we compute their predictive likelihoods at different time periods of the sample. This requires re-estimating the models for different subsets of data. More specifically, we fix seven dates at five year intervals, $\{T_i\}_{i=1}^{7} =$ { 1980:4, 1985:4, 1990:4, 1995:4, 2000:4, 2005:4, 2010:4 }, and re-estimate the models for each dataset $Y_{T_i} = \{y_t\}_{t=1}^{T_i}$ to obtain marginal likelihood values $P(Y_{T_i}|A_j)$, for $i = 1, .., 7$ datasets, and $j = 1, ..., 5$ models. Predictive likelihoods, reported in Table 3 are the log of ratios $P(Y_{T_{i+1}}|A_j)/P(Y_{T_i}|A_j)$ for $i = 1, 2, ..., 6$. Each row in the table reports the results for one of the five models. The second column shows the log of marginal likelihood (ML) in the baseline model over the shortest sample of 101 observations $(1955 : 2 - 1980 : 4)$. Columns 3-6 present the log of predictive likelihood (PL) over the five year intervals, reported as average

[21] Cantore, Levine, and Melina (2014) compare a model similar to our baseline model with one featuring deep habits and claim that modeling deep habits substantially improves the fit of the model. When we estimate the models disregarding the data on taxes, like Cantore, Levine, and Melina (2014), we find that the model with deep habits outperforms the baseline model in terms of model fit, reproducing their result. This finding suggests that the deep habits model has difficulties explaining the dynamics of taxes. Moreover, this also suggests that not accounting for the public financing method or the actual data on taxes may have serious consequences for the results of estimation.

quarterly values. All numbers assume parameter p in the Geweke's estimator is 0.5. The results in the table reveal that even though the deep habits model underperforms in terms of data fit for the overall sample, in the latest 5-year episode, between 2005:4 and 2010:4, this model has the largest predictive likelihood. In different subperiods, other models were the winners in terms of predictive likelihood. For example, during 2000-2005, the model with government spending in the utility, and in the period of 1995-2000, the rule-of-thumb consumer model has the largest predictive likelihood of all models. However, the differences in quarterly predictive likelihoods are not substantial across all models and all 5 year periods we consider. More importantly, we find that the baseline model consistently demonstrates the best data fit than any other model starting at least in the 1980s. Also, the model with deep habits outperforms the rule-of-thumb model for the shortest time interval, however it is also true that in this case all models with specific features are much closer in terms of fit. Consistent with the conclusion of the marginal likelihood analysis, the predictive likelihood evaluation reveal that none of the model specific propagation mechanisms strictly dominate the baseline model in predictive ability, and model fit.

One known criticism of the model comparison using marginal likelihoods is that they can be significantly affected by the choice of parameter prior distributions. While we keep the prior distributions the same for all common model parameters, the prior distributions of model-specific parameters may still be influencing the marginal likelihood of some models in a negative way. An alternative test of model fit based on information criterion is a way to measure the model fit without having to rely on prior distributions in a direct way.[22] While the information criterion based tests are not a Bayesian test, there are modifications that are most useful for making inference in Bayesian estimation. We obtain alternative measures of model fit using the Deviance Information Criterion (DIC) and the Widely Applicable Information criterion (WAIC). Both methods confirm the finding of the marginal likelihood based tests that the baseline model outperforms the other models. The information criterion values for the other four models are weakly equivalent to the rankings implied by marginal likelihoods.[23]

To summarize, the results of the model comparison exercise have the following implication: If a researcher is looking for a model that is quantitatively good at describing the data, then introducing any of the considered transmission mechanisms for government spending may be unnecessary.

[22] Priors may still have influence on parameter estimates, but this criterion does not use priors directly. The sensitivity of estimates to priors is evaluated below in Section 4.4.

[23] The details about the tests and their outcomes are summarized in the Technical Appendix.

4.2 Parameter Estimates

Tables 4 and 5 report the estimated parameters in the five models. The estimates are obtained as mean values over $700,000$ out of 1 million elements of the Markov chain generated using the random walk Metropolis-Hastings algorithm. The proposal distribution is multivariate normal with the variance-covariance matrix $c\Sigma$, where Σ is determined as the inverse of the numerical Hessian evaluated at the starting element for the Markov chain, and $c > 0$ is a parameter that is adjusted to achieve the acceptance rate in the range between 22 and 40 percent, as suggested in Robert and Casella (2005).[24] We relied on a visual evaluation of the trace and cumulative sum (CUSUM) plots to verify that Markov chains are stationary and convergent.

Table 4 documents the estimates of model specific parameters and common parameters, besides the parameters of the shock processes. Although the models have different sets of model specific parameters, habit formation in private consumption is present in all the models. However, this parameter has a slightly different meaning across the models because of the unique specifications of the consumption measure X_t entering utility. Consumption habit parameters for the deep habits model (b^c) and h^c across all the other models is relatively similar and estimated to be close to 0.7 in all cases, which is well within the range reported in the literature.

While the degree of deep habit for private consumption is considerably larger than for public consumption ($b^g = 0.57$), the stock of habit for public consumption depreciates at a slower pace ($\rho^c = 0.07$ and $\rho^{gg} = 0.46$). Model specific parameter for the rule-of-thumb model, which determines the share of population living hand-to-mouth given by λ, is estimated to be 0.1. This number is relatively small (Cogan, Cwik, Taylor, and Wieland (2010) find $\lambda = 0.29$, and using the European data, Forni, Monteforte, and Sessa (2009) estimate $\lambda = 0.34$, while Coenen and Straub (2005) report $\lambda = 0.246$). The estimates for the model with government spending in utility are ν and ϕ. The elasticity of substitution between public and private consumption in the model where government spending enters utility, is $\nu = 1.5$, which is larger than $\nu = 0.3$ estimated in Bouakez and Rebei (2007), and implies less complementarity between government spending and private consumption than in their model. Parameter ϕ has the posterior mean of 0.97 indicating that private consumption is valued by individuals much more than public goods.[25] The mean of the specific parameter in the model with productive government spending, α_G, is estimated at 0.01. This value is close to but smaller than the one calibrated in the study by Baxter and King (1993) (they

[24]The starting element is determined as the mean value of the last $500,000$ draws of another (1 million elements long) Markov chain, the starting element of which coincides with the mean of the prior distribution.

[25]Bouakez and Rebei (2007) calibrate this parameter at 0.8.

use $\alpha_G = 0.05$).

The rest of Table 4 presents the estimates of the common model parameters. While there is some variation, the parameters are generally consistent across the models. The estimate of the price rigidity parameter varies to some degree across the models, and α_p is in the range of 26 to 42. With the exception of the deep habits and rule-of-thumb consumer models, the remaining models demonstrate the wage rigidity parameter above 100. Investment adjustment costs parameter is uniformly greater than 1 in all of the models. It is worth noting that the parameter estimates for the baseline model and the model where spending is productive and utility-enhancing have rather similar estimates across the three. For all the models the parameter σ is estimated to be close to 0.5, which implies an intertemporal elasticity of substitution greater than 1. Parameters of the monetary policy rule in all models imply that the rule is inertial, with α_R estimated in the range of 0.7. The response of the policy interest rate to inflation is moderate, with α_π estimated around 0.3. The estimates imply that the long term response of interest rates to inflation, $\alpha_\pi/(1-\alpha_R)$, is between 1.1 and 1.5 in all models. The response of the interest rate to output growth, measured by α_Y is small and fairly consistent across the models. Finally, the estimates of the tax rule are very similar across all the models - the tax rule is highly inertial (α_τ is approximately 0.95), with a small response to the output gap, where a one percent increase in output gap implies approximately 0.5 basis point increase in the income tax rate.

Table 5 reports the estimates of autocorrelation and standard deviations of the shock processes. The estimates for the government spending, investment specific, and monetary policy shocks are consistent across the models. Consistent with other studies, we find that the government spending process is highly persistent at 0.97. The autocorrelation of the investment specific shock lies within the range of 0.1 - 0.3, and the standard deviation is approximately 0.04. The tax rule estimates are also consistent across all models, where the average tax rate is significantly persistent and has a positive response to output deviations in the range of 0.004 and 0.006 across all models. On the other hand, the models provide quite different estimates for the neutral technology and preference processes. It is important to understand however, that neither model can perfectly describe the properties of the data. When an estimated model is missing an internal mechanism to replicate some properties of the data, such as autocorrelations and volatilities, then the estimates of shock processes will be adjusted to replicate observed correlations in the data.

4.3 Impulse Response Functions

There has been a lot of debate in the literature about the effect of government spending on private consumption. The models we investigate in this paper were all developed as a channel to allow consumption to rise in response to an unexpected increase in government spending observed in many empirical structural VAR models. The literature however, has still not come to an agreement on this issue. While some authors report evidence favoring the positive response (see Ravn, Schmitt-Grohé, and Uribe (2007), Bouakez and Rebei (2007), Zubairy (2014b)), others fail to find it in their estimated models (see for example Leeper, Plante, and Traum (2010), Coenen and Straub (2005)). We address this debate by comparing responses of consumption to the government spending shock across the estimated models.

Figure 3 plots the impulse response of consumption to a 1 percent increase in government spending in the five estimated models, shown as percentage deviations from trend, with quarters along the horizontal axis.[26] Strikingly, none of the five models under consideration predicts a positive response of consumption to the shock. In fact, even on impact the consumption response is negative and significantly different from zero among all of the models under consideration.

Note that all five models have the potential to predict a positive consumption response to a government spending shock. To verify this, we conduct a prior predictive analysis of the consumption response, in the same spirit as Leeper, Traum, and Walker (2011). More precisely, we draw a random sample of 1000 elements for the model parameters from their respective prior distributions, and then compute the impulse response functions. Figure 5 shows the distribution of the consumption response to a positive 1 percent government spending shock across the five models. The figure reports the median as well as the 5th and 95th quantile of the prior distribution as the dashed lines. Figure 5 also shows the posterior impulse responses, which are replicated from Figure 3. There are a few things to note. Firstly, it is clear that the priors imply a range of the consumption response that spans both positive and negative values on impact across all five models, which indicates that the estimated negative consumption responses are induced by the data, rather than by prior distributions.[27] Secondly, the posterior distributions of the impulse responses are quite similar in all estimated variants although the priors can be quite different across the various

[26]The responses to a 1 percent shock are obtained by normalization. Namely, we calculate impulse responses to a shock of the size equal to one standard deviation, which is drawn from the posterior distribution, and then normalize all responses by the median estimate for standard deviation before presenting the responses in percentages.

[27]In light of the possible identification issues described above, we also apply the Müller's criterion to evaluate how the effect of priors may influence the immediate response of consumption. We find that the estimated impact response of consumption is robust to the choice of prior distribution in all models. See the Technical Appendix for more details.

models.

It is possible to relate the negative consumption response to the fact that in all these estimated models, the posterior distribution of intertemporal elasticity of substitution for consumption, σ, turns out to be smaller than one. When $\sigma < 1$, consumption and leisure are complements, meaning that $U_{12} > 0$. This means that an equilibrium increase in labor (and drop in leisure) is associated with the decreasing marginal utility of consumption, providing incentives for the households to reduce consumption. In the case of $\sigma > 1$ instead, consumption and leisure are substitutes in the sense that $U_{12} < 0$. Therefore, the opposite is true: An increase in hours worked h raises marginal utility of consumption, making it more desirable for households to raise consumption.

Given the parameter estimates, it should be expected that the responses of consumption in the four models are similar that of the baseline model. The closest to the baseline model is the model with productive public expenditures, and the estimated parameter (α_G) is close to zero. In the model with public spending in the utility function, we find that public and private spending are closer to substitutes rather than complements, and the weight on public spending $(1-\phi)$ is close to zero. The rule-of-thumb consumer models also predict a negative response of consumption. The estimated share of liquidity constrained household in the rule-of-thumb model, λ at approximately 10 percent of the population, is not large enough to ensure positive response of aggregate consumption. In the deep habits model, the degree of deep habit in public consumption is not nearly as big as that in Zubairy (2014b).

The negative response of consumption to a spending shock found across all five models is contrary to some other studies. While a detailed analysis of the source of differences in the consumption response from past studies is beyond the scope of the paper, we believe they can be attributed to differences in model specifics or the choice of observables. For example, in the estimated model of Zubairy (2014b) featuring deep habits, consumption increases in response to the government spending shock, but is estimated using a different set of observables.[28] The differences may also be due to the poor identification of the parameter of deep habit in public spending parameter, b^g, an important parameter in driving the response of consumption. In addition, Fève, Matheron, and Sahuc (2013) show that if the endogenous component of government policy is explicitly accounted for in the spending process, then a higher degree of Edgeworth complementarity is needed to match the correlation between private consumption and government spending, and higher complementarity may help reproduce a positive response of consumption to a spending shock. Also, Bouakez and Rebei

[28] We consider labor variables, including hours and real wage as observables, while Zubairy (2014b) does not, focusing instead on other fiscal variables. Also, we model stochastic growth in technology shocks, while that paper assumes stationary shock processes.

(2007) find a positive consumption response in a model featuring government spending in the utility function, however notably they use a very small set of observables (only four series) and it does not include government spending. However, they calibrate rather than estimate some important parameters of the utility function.

While all the five distinct models considered predict a fall in consumption in response to a positive government spending shock, we also consider one additional model variant, namely a specification where we allow for all the features in the five variants at the same time, i.e. we allow for deep habits, rule-of-thumb consumers, productive government spending, and non-separability between private consumption and government spending, all simultaneously.[29] This model, obviously, has a prior mode path for private consumption to a spending shock which is much more strongly positive. We find that the estimated "nested" model indeed produces a positive response of consumption, and the positive response of consumption is largely due to the government spending in the utility function specification. However, based on marginal likelihood, this model fares rather badly relative to the other models in terms of overall data fit.[30]

There has also been an increased interest in the literature regarding the size of the government spending multiplier. We contribute to this debate in Figure 4, which plots output responses to a 1 percent increase in government spending across the five models. The responses are shown in percentage deviations from the non-stochastic trend, with quarters along the horizontal axis. We find that the response of output is similar for the baseline model, and the models featuring productive and utility-enhancing government spending and the model featuring deep habits. In these models, output rises close to 0.17 percent in response to a 1 percent government spending shock. Given that the steady state share of government spending was fixed at 0.2, this translates into a government spending multiplier of less than one on impact, at approximately 0.85 in these three models.[31] The response of output is slightly larger in the model featuring rule-of-thumb consumers, which implies that on impact, the government spending multiplier for output is approximately 1 in this model. Therefore, we conclude that most of the models agree on the size of the impact government spending multiplier of output being less than 1. In addition, none of them predict a multiplier that is substantially larger than 1.

[29]This is not strictly a nested model, since we have external habit formation as in the deep habit model and not internal habit formation as considered in the other models.

[30]Please refer to the Technical Appendix for more details on this model.

[31]The multiplier is computed as the impact response of output divided by the steady state ratio of public spending to output, which is 0.2.

4.4 Identification of Parameters

In an ideal environment where the data contains enough information about the parameter of interest, the choice of the prior, if proper, should not influence the estimation results. However, in practice, some parameters of interest may be poorly identified and the choice of prior in this case may substantially influence the posterior distribution. While the absence of identification is not a severe problem for a Bayesian economist, we believe it is important to understand identifiability properties of the estimated parameters, as well as to understand the implications for model comparison and impulse response analysis.

We evaluate the prior and posterior distributions, and note that for most parameters, the prior distributions are wide, and posterior distributions are different from the priors.[32] This is, however, not true for all parameters. For instance, the posterior distributions for two parameters, μ_p and μ_w seem to be entirely influenced by their priors.

In order to gauge the ability of our empirical strategy to identify the model parameters, we perform two formal identification tests. Firstly, we conduct the test proposed by Iskrev (2010), which essentially checks whether the derivatives of the predicted autocovariogram of the observable variables with respect to the estimated parameters has rank equal to the number of estimated parameters. Applying this testing procedure reveals that the pairs of parameters μ_p, σ_p, and μ_w, σ_w are not identifiable in all the models. It seems that a relatively strong prior on the μ parameters allows us to estimate the variance parameters, σ_p and σ_w. We also find that the derivative of the vector of predicted autocovariogram of the vector of observables, with respect to a vector of estimated parameters has full rank when we exclude these four parameters and evaluate it at the posterior mean estimate. Therefore, the results suggest that the other parameters are identifiable in the neighborhood of our estimates.

Additionally, we carry out the test proposed by Müller (2012), which allows us to evaluate identifiability of the estimate by estimating the sensitivity of the posterior mean to the mean of prior distribution. In one-dimensional estimation, this sensitivity is measured by the ratio of posterior variance to prior variance. In estimation with multidimensional parameter of interest, the prior sensitivity measure is derived similarly from covariances of the prior and posterior distribution, which takes into account the effects of priors and posterior of some estimated parameters on posterior means of others. The results from this test are in line with the finding of the Iskrev's test, suggesting that the influence of the prior distribution may be important for the parameters μ_p, μ_w, as well as for the parameter determining the slope of the depreciation function, δ_2/δ_1, the parameter of deep habit in public spending b^g, and the utility parameter μ, which determines the degree of complementarity between

[32]These are shown for each model in the Technical Appendix.

private and public consumption.

Therefore, these identification tests point towards problems in identification of some parameters, especially those related to the price and wage markup shock processes. While we do not fix these parameters, one may think of imposing identification by the choice of priors for μ_p and μ_w, which are close to the estimates in Smets and Wouters (2007). More importantly, these identification tests suggest that the estimated parameters in each of the models considered, that play an important role in the transmission mechanism of the spending shocks are identifiable, except likely in the deep habits and the public spending in the utility models.

Lastly, it is important to evaluate whether the prior distributions influence the resulting impulse responses government spending shock. Namely, we use Müller's prior informativeness criterion to evaluate the effect of priors on the contemporaneous response of consumption and output to the government spending shock. We find that this effect is overall small.[33]

4.5 Moments

Bayesian estimation, being a full information approach, effectively is attempting to induce all moments and cross-correlations from the data into an estimated model. To shed some light on the relative model fit, it is therefore useful to compare moments and autocorrelations implied by the estimated models to those implied by the data.[34] Tables 6 and 7 report unconditional moments and autocorrelations. We obtain these statistics for the models by simulating datasets of 221 observations. The procedure is repeated 1000 times to evaluate the 5^{th} and 95^{th} quantiles of the posterior moment distributions, presented in square brackets, and to report the estimates as the median values of these distributions. All simulated data are analogues of the observable variables in the data. In particular, the statistics for investment, consumption, and the wage rate are those of growth rates of model variables, government spending statistics is government spending relative to the previous period output level, while inflation and interest rate are annualized in both the data and the model. The last column displays the statistics for the vector of observable variables used in estimation.

Table 6 focuses on standard deviations, both in absolute terms and relative to the standard deviation of output growth for the five models and in the data.[35] While the models

[33]These results are available in Technical Appendix.

[34]This exercise may also be important as an additional check of model fit in light of the fact that marginal likelihoods may be affected by prior distributions.

[35]Even though output is not directly a part of estimation, we use output growth rather than consumption as a basis for comparison, because it is common in the literature, which simplifies cross study comparisons. The measure of output in the data is GDP, and the model measure of output is production function less the costs of price changes.

over-predict the volatilities of almost all variables, besides the fiscal variables, we find that they do better at matching the standard deviations in relative terms, matching especially well the relative volatility of investment, hours and wages in all the models. As in the data, all models report that investment is more volatile, and consumption is less volatile, in terms of the growth rates. The models with deep habits and rule of thumb consumers demonstrate statistically more volatile consumption relative to output than what is observed in the data.[36] This inability to match the volatility of consumption growth could potentially explain the poorer fit of these models as compared to the rest of the models. However, these two models provide statistically better match to the relative standard deviation of inflation than the other models considered.

Table 7 presents the serial correlations and correlations with output growth implied by the models and the data. The top part of the table reports serial correlation for the observed variables. One can see that all the models do especially well at matching the persistence in variables that are highly correlated in the data - government spending variable, tax rate, hours, inflation and the interest rate, but fail to match autocorrelations of variables with small and moderate autocorrelation coefficients, particularly investment and the wage rate. All models match autocorrelation in consumption in a statistical sense, however the rule of thumb consumers model reproduces this statistics exceptionally well, which again must be due the fact that consumption of constrained households is proportional to their income.

The lower part of Table 7 displays the correlation between observable variables and output growth.[37] In all models, relative correlation of government spending variable, tax rate, investment and consumption matches those from the data reasonably well. Interestingly, the model with rule of thumb consumers is again the best in matching the relative correlation of consumption. All models, however, demonstrate poor fit in relative correlations of hours, wages, inflation and the rate of interest. The poor overall performance of the rule-of-thumb model is apparent in its prediction of the correlation between inflation and output growth, which has the largest discrepancy among all the models considered and is predicted to be strongly positive, when the data suggests a negative correlation between the two. Out of all the models, the model featuring deep habits notably comes closest to matching the negative correlation between inflation and output growth. This can be related to the fact that in the deep habits model, firms take into account future expected demand relative to the current demand when setting prices, and thus the Phillips curve has both current and expected

[36]The increased consumption volatility in the rule of thumb model could be explained by the presence of liquidity constrained consumers who are not able to insure against risk.

[37]Output in the data moments is measured by GDP, while output in model implied moments is the model consistent measure of output.

future demand in the expression.[38] All the models predict a negligible correlation between government spending variable and output growth, which is consistent with the data.

Since output is not an observable variable, we also looked at the correlation of the observable variables with consumption growth, which is an observable in the estimation.[39] In that case too, the models match the moments reasonable well. Specifically, the unconditional correlation between government spending variable and consumption growth in the data is also insignificant, at -0.01. All the estimated models successfully predict an insignificant negative correlation between the two. It is possible that the negative response of consumption to the government spending shock in estimated models is the result of an attempt to reproduce this negative correlation.

Table 8 reports the contribution of the government spending shock to the overall volatility of macroeconomic variables implied by each model. Each column in the table shows model implied standard deviation of a variable when government spending is the only source of uncertainty, relative to the unconditional standard deviation of this variable assuming all sources of uncertainly are present, in percentages. Consistent with other studies, in all models the government spending shock explains only a moderate fraction of volatility for most variables. Specifically, the contribution of the government spending shock to consumption does not exceed 10 percent, which points to the limited importance of government spending shocks in generating consumption fluctuations. Moreover, the prevalence of other shocks could explain the positive correlation of consumption and output growth even when the response of consumption to the expansionary government spending shock is negative in all models.

4.6 Relaxing the DSGE Model Restrictions: Implications of DSGE-VAR Estimation

While DSGE models are extremely popular in quantitative research, concerns have been raised that these models impose restrictions that are too tight to successfully replicate the dynamics of macroeconomic variables (see Del Negro, Schorfheide, Smets, and Wouters (2005)). Del Negro and Schorfheide (2004) suggest that tight restrictions can be relaxed in estimation of a DSGE-VAR model. In this section, we evaluate the robustness of our conclusions by embedding the transmission mechanisms in a DSGE-VAR framework following Del Negro and Schorfheide (2004). A DSGE-VAR is a VAR model where a DSGE model implies some prior distributions for coefficients and the covariance matrix of innovations. The priors are

[38] Zubairy (2014a) provides a more detailed discussion of the Phillips curve under deep habits.

[39] These correlations are shown in the Technical Appendix.

incorporated by augmenting the actual data set with the data generated by a theoretical model. We choose to estimate a VAR with four lags, as Del Negro, Schorfheide, Smets, and Wouters (2005) show that four lags is sufficient for the VAR model to approximate the true model dynamics very well. Suppose Φ_i, for $i = 1, ..., 4$ are estimates of matrix coefficients next to lagged values of observable variables in the reduced form VAR(4) model, and Σ_u is the estimate of the reduced form innovations' covariance matrix. The dynamics of the DSGE model can also be represented approximately by a reduced form VAR with four lags, with the matrix coefficients $\Phi_i^*(\theta)$, for $i = 1, ..., 4$ being the nonlinear functions of parameter vector θ, and u_t is the vector of reduced form innovations, with the implied covariance $\Sigma_u^*(\theta)$, which is also a nonlinear function of θ. The prior distribution for θ implies the priors for matrices Φ_i^*, $i = 0, 1, ..., 4$, and Σ_u^*, while these priors, combined with the likelihood function, provide posterior distributions for the VAR coefficients, the variance of the reduced form innovations, and the parameter vector θ.[40]

Table 9 reports the resulting marginal likelihood across the five DSGE-VAR models, which is calculated using the strategy outlined in Section 3.2. The table reveals that the ranking of the transmission mechanisms embedded in the DSGE-VAR framework is preserved for all the models except the productive government spending model. However, while the marginal likelihood of this model exceeds that of the baseline model, the log Bayes factor is no greater than 0.9, which is not indicative of a better model performance.

Comparing the results in Tables 2 and 9, one may notice that the marginal likelihood values are substantially larger for the DSGE-VAR models, which means that these models outperform DSGE models in terms of the data fit. This conclusion is in line with Del Negro, Schorfheide, Smets, and Wouters (2005), and indicates that the DSGE frameworks are misspecified to a certain degree. This is also confirmed by the relatively small marginal likelihood maximizing value for λ, the optimal weight for the DSGE model, which we find to be 0.4. However, while all the models reveal some degree of misspecification, it is important to note that neither the improvement in fit of the DSGE-VAR over the DSGE model, nor the small value of λ necessarily imply that the DSGE model are misspecified to the extent that their predictions cannot be trusted.[41]

To further assess the extent of model misspecification, we calculate the impulse responses implied by the estimated DSGE-VAR models, and compare them with the responses from the theoretical models. Because the DSGE-VAR models are formulated in reduced form, producing impulse responses requires identification of the government spending shocks. We

[40] More details on calculating the posterior distributions can be found in Del Negro and Schorfheide (2004).

[41] The potential problem of model misspecification is a more general issue we are not trying to resolve in this paper. Even if the modeling framework is potentially misspecified, it is uniformly misspecified across all models and therefore is unlikely to influence the outcome of our comparison study.

follow the identification procedure outlined in Del Negro and Schorfheide (2004). In particular, we assume that the reduced form innovations of the DSGE-VAR model, u_t, and true shocks ϵ_t are linearly related as follows:

$$u_t = \Sigma\Omega^* \epsilon_t,$$

where Σ is a lower triangular matrix, and Ω^* is an orthonormal matrix. Σ can be obtained as the Cholesky decomposition of Σ_u, and Ω^* is calculated from the theoretical model by decomposing the matrix of immediate responses of observables to shocks into a lower triangular matrix Σ^* and the orthonormal matrix Ω^*[42]

$$\frac{\partial y_t}{\partial \epsilon_t} = \Sigma^* \Omega^*,$$

where y_t is the vector of observables implied by the model. Because government spending is the first variable in the observable vector, the Cholesky triangularization of the variance of the reduced form innovations implicitly assumes that unexpected changes in government spending can only be affected by the government spending shock, as in Blanchard and Perotti (2002).

We find that, with one exception, the impulse responses to the government spending shock for all the theoretical models resemble those of the DSGE-VAR model reasonably well, indicating that the degree of model misspecification is probably only moderate.[43] The exception is the consumption response, which is positive in the DSGE-VAR models. The consumption responses for all the five models are presented in Figure 6, together with the 5^{th} and 95^{th} quantiles.[44] The dramatic difference in short-run consumption responses with the DSGE responses points to a substantial mismatch between the theoretical models and the data.

In order to better understand the source of discrepancy in the consumption responses, it is important to notice that the DSGE-VAR model responses resemble those from the SVAR model quite closely.[45] Therefore, the main identifying assumption under Blanchard

[42]The use of such a rotation matrix ensures that in the absence of misspecification, the impulse responses to all shocks in the DSGE and the DSGE-VAR models coincide.

[43]The plots of DSGE-VAR implied model responses are available in the Technical Appendix.

[44]We use the posterior parameter distributions to evaluate the mean impulse responses and the confidence intervals in the figure.

[45]This similarity reveals the prevailing importance of the data over the model, but also it is due to the fact that our first observable variable directly measures the government spending shock, and therefore the (Cholesky identified) innovation of the first equation in the DSGE-VAR model is equivalent to the government spending shock in the model (the first column of the rotation matrix Ω^* is a unity vector). The SVAR model responses are reproduced in the Technical Appendix.

and Perotti (2002) that government spending is pre-determined relative to output heavily influences the DSGE-VAR model responses. Note that the government spending process in our DSGE model, given by Equation (10) is also consistent with this assumption. However, Fève, Matheron, and Sahuc (2013) suggest that this kind of government spending process may be a source of misspecification for the DSGE model. Notably, they show that allowing for contemporaneous response of output has implications for model fit, and more importantly, omitting it results in a downward bias of consumption response to a spending shock in an estimated DSGE model. Similarly, Kormilitsina (2017) tests the validity of Blanchard and Perotti (2002) identifying restrictions in a DSGE model by evaluating whether government spending responds to realizations of macroeconomic shocks in the same quarter.[46] She finds strong evidence that government spending responds to current realizations of technology shocks. Also, the paper shows that the failure to account for endogeneity of government spending with respect to output results in an upward bias of the consumption response in an SVAR model.

While these papers provide directions for reconciling the DSGE and DSGE-VAR responses to a government spending shock for consumption, our choice of government spending process was driven by the most commonly used specification in the literature, and one that is consistent with Blanchard and Perotti (2002), as described in more detail in Section 2.1.3. Moreover, we do not find evidence that allowing for contemporaneous response to output in our spending process results in a better model fit.[47]

To summarize, we find that the estimation of DSGE-VAR models confirms the rankings of the propagation mechanisms. However, the consumption effect of government spending shocks is positive in the DSGE-VAR frameworks, contrary to the negative response in DSGE models, which seems to be influenced by the Blanchard and Perotti (2002) shock identification. For a researcher with strong prior beliefs in the validity of this identification, this suggests that the DSGE models substantially misrepresent the dynamics of consumption in light of the government spending shock, and none of the transmission mechanisms we study helps to resolve this problem.[48] On the other hand, there is still disagreement in the fiscal literature about the correct methodology to identify shocks to government spending.[49]

[46] Kormilitsina (2017) adopts a standard DSGE model with informational subperiods which allows for a timing structure where pass through of information about the state of the economy can be controlled.

[47] In order to allow government spending to respond to output contemporaneously, the simplest modification in our framework is to modify the definition of the government spending shock in Equation (10) by substituting G_t/Y_{t-1} with G_t/Y_t. Model comparison for our baseline model reveals that the data fits the data better with the original specification with G_t/Y_{t-1} than with G_t/Y_t.

[48] Another potential interpretation of the results is that a model with pre-determined government spending, as dictated by Equation (10), may be misspecified, as discussed above, based on Fève, Matheron, and Sahuc (2013) and Kormilitsina (2017).

[49] Besides the timing restrictions in a SVAR setting (Blanchard and Perotti (2002)), identification strategies

5 Conclusion

In this paper, we rely on Bayesian estimation to quantitatively investigate five distinct models of the government spending shock. We find that the basic new-Keynesian model with nominal frictions, where government spending represents a mere waste of economy's resources, and utility is nonseparable in consumption and leisure, fits the data just as well as the alternatives, if not better in some cases. The remaining four models feature propagation mechanisms for the government spending shock that were originally introduced in an attempt to explain a positive correlation between private and public consumption conditional on a spending shock, a finding often documented in empirical research. Namely, the models in this study incorporate deep habits in consumption, rule-of-thumb consumers, government spending directly influencing utility of economic agents, and the idea of productive public capital. However, we do not find support for this co-movement hypothesis in any of the models we consider. In particular, all the estimated DSGE models predict a drop in consumption as a response to an unexpected rise in government spending.

When we apply the DSGE-VAR methodology for each model, the ranking of the various propagation mechanisms is preserved. However, in departure from the DSGE models, the response of consumption to a government spending shock is positive in the DSGE-VAR framework under all the various models. This points towards issues of potential misspecification in the DSGE models. Most notably given the inconsistency between the consumption response, if one strongly believes in the Blanchard and Perotti (2002) identification scheme which finds a positive response of consumption to a spending shock, it highlights the specific shortcoming of all commonly used fiscal DSGE models. This is of particular note since our study was motivated by the DSGE literature trying to incorporate various transmission channels to reconcile the positive response of consumption to a spending shock found in SVARs. However, this disagreement between the DSGE and DSGE-VAR models also raises questions about the validity of assuming pre-determined government spending in fiscal DSGE models in order to be consistent with Blanchard and Perotti (2002) identification. This is further supported by the recent works of Fève, Matheron, and Sahuc (2013) and Kormilitsina (2017).[50]

include narrative evidence on military buildups (Ramey (2011)) and sign restrictions (Mountford and Uhlig (2009)), to name a few.

[50]Notably, both Fève, Matheron, and Sahuc (2013) and Kormilitsina (2017) provide evidence that government spending responds to output in the same quarter and show that assuming pre-determined government spending can bias the consumption response to a spending shock.

References

Altig, D., L. Christiano, M. Eichenbaum, and J. Linde (2011): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," *Review of Economic Dynamics*, 14(2), 225–247.

Ambler, S., and A. Paquet (1996): "Fiscal Spending Shocks, Endogenous Government Spending, and Real Business Cycles," *Journal of Economic Dynamics and Control*, 20(1-3), 237–256.

Baxter, M., and R. G. King (1993): "Fiscal Policy in General Equilibrium," *American Economic Review*, 83(3), 315–34.

Bilbiie, F. O. (2009): "Nonseparable Preferences, Fiscal Policy Puzzles, and Inferior Goods," *Journal of Money, Credit and Banking*, 41(2-3), 443–450.

Bils, M. (1987): "The Cyclical Behavior of Marginal Cost and Price," *American Economic Review*, 77(5), 838–55.

Blanchard, O., and R. Perotti (2002): "An Empirical Characterization Of The Dynamic Effects Of Changes In Government Spending And Taxes On Output," *The Quarterly Journal of Economics*, 117(4), 1329–1368.

Bouakez, H., and N. Rebei (2007): "Why does Private Consumption Rise After a Government Spending Shock?," *Canadian Journal of Economics*, 40(3), 954–979.

Cantore, C., P. Levine, and G. Melina (2014): "Deep versus Superficial Habit: It's all in the Persistence," School of Economics Discussion Papers 0714, School of Economics, University of Surrey.

Coenen, G., and R. Straub (2005): "Does Government Spending Crowd in Private Consumption? Theory and Empirical Evidence for the Euro Area," *International Finance*, 8(3), 435–470.

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.

Colciago, A. (2011): "Rule-of-Thumb Consumers Meet Sticky Wages," *Journal of Money, Credit and Banking*, 43, 325–353.

Del Negro, M., and F. Schorfheide (2004): "Priors from General Equilibrium Models for VARS," *International Economic Review*, 45(2), 643–673.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2005): "On the fit and forecasting performance of New-Keynesian models," Working Paper Series 0491, European Central Bank.

Fatas, 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.

Fève, P., J. Matheron, and J.-G. Sahuc (2013): "A Pitfall with Estimated DSGE-Based Government Spending Multipliers," *American Economic Journal: Macroeconomics*, 5(4), 141–78.

Fisher, J. (2003): "Technology Shocks Matter," FRB of Chicago w.p #2002-14, December.

Fisher, J., and R. Peters (2010): "Using Stock Returns to Identify Government Spending Shocks," *Economic Journal*, 120(544), 414–436.

Forni, L., L. Monteforte, and L. Sessa (2009): "The General Equilibrium Effects of Fiscal Policy: Estimates for the Euro Area," *Journal of Public Economics*, 93(3-4), 559–585.

Furlanetto, F. (2011): "Fiscal Stimulus and the Role of Wage Rigidity," *Journal of Economic Dynamics and Control*, 35(4), 512–527.

Galí, J., J. D. Lopez-Salido, and J. Vallés (2007): "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association*, 5(1), 227–270.

Ganelli, G., and J. Tervala (2009): "Can Government Spending Increase Private Consumption? The Role of Complementarity," *Economics Letters*, 103(1), 5–7.

Gelfand, A. E., and D. K. Dey (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society. Series B (Methodological)*, 56(3), 501–514.

Geweke, J. (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18(1), 1–73.

Geweke, J., and G. Amisano (2010): "Comparing and evaluating Bayesian predictive distributions of asset returns," *International Journal of Forecasting*, 26(2), 216–230.

Iskrev, N. (2010): "Local identification in DSGE models," *Journal of Monetary Economics*, 57(2), 189–202.

Jacob, P. (2013): "Deep Habits, Price Rigidities and the Consumption response to Government spending," Reserve Bank of New Zealand Discussion Paper Series DP2013/03.

Jeffreys, H. (1961): *The Theory of Probability*. Oxford University Press.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988): "Production, Growth and Business Cycles : The Basic Neoclassical Model," *Journal of Monetary Economics*, 21(2-3), 195–232.

Kormilitsina, A. (2013): "Solving Rational Expectations Models with Informational Subperiods: A Perturbation Approach," *Computational Economics*, 41(4), 525–555.

——— (2017): "Is Government Spending Predetermined? A Test of Identification for Fiscal Policy Shocks," Discussion paper.

Leeper, E. M., M. Plante, and N. Traum (2010): "Dynamics of Fiscal Financing in the United States," *Journal of Econometrics*, 156(2), 304–321.

Leeper, E. M., N. Traum, and T. B. Walker (2011): "Clearing Up the Fiscal Multiplier Morass," NBER Working Paper 17444.

Linnemann, L. (2006): "The Effect of Government Spending on Private Consumption: A Puzzle?," *Journal of Money, Credit and Banking*, 38(7), 1715–1735.

Linnemann, L., and A. Schabert (2003): "Fiscal Policy in the New Neoclassical Synthesis," *Journal of Money, Credit and Banking*, 35(6), 911–29.

——— (2004): "Can Fiscal Spending Stimulate Private Consumption?," *Economics Letters*, 82(2), 173–179.

——— (2006): "Productive Government Expenditure In Monetary Business Cycle Models," *Scottish Journal of Political Economy*, 53(1), 28–46.

Monacelli, T., and R. Perotti (2008): "Fiscal Policy, Wealth Effects, and Markups," NBER Working Papers 14584, National Bureau of Economic Research, Inc.

Mountford, A., and H. Uhlig (2009): "What are the Effects of Fiscal Policy Shocks?," *Journal of Applied Econometrics*, 24(6), 960–992.

Müller, U. K. (2012): "Measuring prior sensitivity and prior informativeness in large Bayesian models," *Journal of Monetary Economics*, 59(6), 581–597.

Perotti, R. (2008): "In Search of the Transmission Mechanism of Fiscal Policy," in *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, pp. 169–226. National Bureau of Economic Research, Inc.

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

Ravn, M., S. Schmitt-Grohé, and M. Uribe (2006): "Deep Habits," *Review of Economic Studies*, 73(1), 195–218.

Ravn, M. O., S. Schmitt-Grohé, and M. Uribe (2007): "Explaining the Effects of Government Spending Shocks on Consumption and the Real Exchange Rate," NBER Working Papers 13328, National Bureau of Economic Research, Inc.

Robert, C. P., and G. Casella (2005): *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

ROTEMBERG, J. J. (1982): "Sticky Prices in the United States," *Journal of Political Economy*, 90(6), 1187–1211.

ROTEMBERG, J. J., AND M. WOODFORD (1991): "Markups and the Business Cycle," in *NBER Macroeconomics Annual 1991, Volume 6*, NBER Chapters, pp. 63–140. National Bureau of Economic Research, Inc.

——— (1999): "The Cyclical Behavior of Prices and Costs," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 16, pp. 1051–1135. Elsevier.

SMETS, F., AND R. WOUTERS (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

TRAUM, N., AND S. S. YANG (2015): "When Does Government Debt Crowd Out Investment?," *Journal of Applied Econometrics*, 30(1), 24–45.

ZUBAIRY, S. (2014a): "Interest Rate Rules And Equilibrium Stability Under Deep Habits," *Macroeconomic Dynamics*, 18(01), 23–40.

——— (2014b): "On Fiscal Multipliers: Estimates From A Medium Scale DSGE Model," *International Economic Review*, 55, 169–195.

6 Tables and Figures

Table 1: Parameter calibration and steady state values

Common parameters		
θ	Production: capital share	0.3
μ_{z^*}	Growth of output	1.0047
μ_{Υ}	Growth of investment specific technology	1.0042
π	Steady state inflation	1.0086
β	Intertemporal discount factor	0.999
δ_0	Depreciation rate	0.025
ϱ	Shadow price of capital	1
u	Steady state rate of capital utilization	1
h	Steady state labor	0.5
s^g	Steady state share of govt. spending in output	0.2
η_p	Prices: elasticity of substitution	6
η_w	Wages: elasticity of substitution	21
τ	Steady state tax rate	0.18
Model with productive public capital		
s^{gc}	Consumption share of public spending	0.8

Table 2: Model Marginal Likelihood

p	BL	DH vs. BL	ROT vs. BL	G in U vs. BL	G in F vs. BL
0.1	5092.4	-34.1	-22.3	-7.0	-4.3
0.5	5094.5	-34.4	-22.3	-7.2	-4.4
0.9	5096.4	-34.9	-22.1	-7.2	-4.1

Notes. Table shows logarithm of marginal likelihood of a model evaluated using Geweke (1999) procedure. The first column is the parameter p in the Geweke estimator that specifies the supplementary p.d.f $f(\theta)$ in Equation (17). The second column shows the marginal likelihood in the baseline model. Columns 3-6 present the log of marginal likelihood of a model relative to the best fitted model, so that negative numbers indicate more poor fit. DH = model with deep habits, ROT = for rule-of-thumb model, G in U = model with government spending in the utility function, G in F = the model with government spending in the production technology, and BL = the baseline model with no specific features.

Table 3: 5-year Predictive Model Likelihoods.

Model	1954-79	1980 -84	1985-90	1991-95	1996-00	2001-05	2006-10
	ML	PL	PL	PL	PL	PL	PL
DH	2205.1	22.3	25.1	24.9	24.1	22.8	23.6
ROT	2201.4	22.3	24.8	25.2	24.5	23.3	23.5
G in U	2205.8	22.8	25.0	25.2	24.2	23.8	23.2
G in F	2208.1	22.8	25.0	25.2	24.2	23.7	23.2
BL	2218.6	22.6	24.9	25.2	24.2	23.7	23.2

Notes. Table shows logarithm of predictive likelihood of models. Each row reports the results for one of the five models. The second column shows the log of marginal likelihood (ML) in the baseline model over the first 101 observations. Columns 3-6 present the log of predictive likelihood (PL) over a 5 year interval, reported as an average value for the quarters. All numbers assume parameter p in the Geweke estimator is 0.5. DH = model with deep habits, ROT = for rule-of-thumb model, G in U = model with government spending in the utility function, G in F = the model with government spending in the production technology, and BL = the baseline model with no specific features.

Table 4: Parameter Estimates: Part I

Parameter	Prior distribution		DH	ROT	G in U	G in F	Baseline
	Type	Mean (st.d.)	Mean (st.d.)	Mean (st.d.)	Mean (st.d.)	Mean (st.d)	Mean (st.d.)
h^c	B	0.5	-	0.689	0.772	0.782	0.774
		(0.2)		(0.040)	(0.024)	(0.024)	(0.024)
b^c	B	0.5	0.744	-	-	-	-
		(0.2)	(0.023)				
ρ^c	B	0.5	0.072	-	-	-	-
		(0.2)	(0.032)				
b^g	B	0.5	0.567	-	-	-	-
		(0.2)	(0.296)				
ρ^{gg}	B	0.5	0.464	-	-	-	-
		(0.2)	(0.167)				
λ	B	0.2	-	0.098	-	-	-
		(0.1)		(0.025)			
α_G	I	0.1	-	-	-	0.012	-
		(0.1)				(0.004)	
ν	G	0.8	-	-	1.497	-	-
		(0.5)			(0.558)		
ϕ	B	0.7	-	-	0.967	-	-
		(0.1)			(0.019)		
α_p	G	20.0	42.000	27.028	31.201	31.247	31.166
		(5.0)	(5.606)	(2.730)	(4.094)	(4.213)	(4.108)
α_w	G	100.0	77.892	83.591	123.329	134.573	123.004
		(30.0)	(19.013)	(13.162)	(17.234)	(22.007)	(17.025)
α_R	B	0.7	0.732	0.732	0.713	0.716	0.712
		(0.2)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
α_π	G	0.5	0.345	0.318	0.338	0.336	0.338
		(0.2)	(0.026)	(0.022)	(0.021)	(0.021)	(0.022)
α_Y	G	0.1	0.044	0.070	0.035	0.034	0.035
		(0.1)	(0.008)	(0.011)	(0.008)	(0.008)	(0.008)
κ	G	1.0	1.643	1.506	1.216	1.243	1.250
		(0.5)	(0.329)	(0.211)	(0.144)	(0.148)	(0.145)
δ_2/δ_1	G	2.0	5.683	3.896	5.588	5.540	5.562
		(0.5)	(0.700)	(0.547)	(0.681)	(0.674)	(0.686)
σ	G	2.0	0.682	0.680	0.438	0.436	0.439
		(0.5)	(0.026)	(0.074)	(0.047)	(0.046)	(0.044)
α_τ	B	0.8	0.949	0.949	0.948	0.947	0.948
		(0.1)	(0.014)	(0.011)	(0.013)	(0.013)	(0.013)
$\alpha_{\tau,y}$	I	0.0	0.005	0.004	0.006	0.006	0.006
		(0.1)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)

Notes. Table shows prior distributions and Bayesian estimates of parameters across different models. Notation in the second columns is as follows: B = beta, G = gamma, I = inverse gamma distributions. Estimates are presented as mean values and standard deviations across the last $700,000$ out of 1 million elements of a Markov chain generated using the Metropolis Hastings algorithm. Kalman filter is used to evaluate the likelihood of the data.

Table 5: Parameter Estimates: Part II

Parameter	Prior distribution		DH	ROT	G in U	G in F	Baseline
	Type	Mean (st.d.)	Mean (st.d.)	Mean (st.d.)	Mean (st.d.)	Mean (st.d)	Mean (st.d.)
ρ_g	B	0.9	0.967	0.969	0.968	0.967	0.968
		(0.1)	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
ρ_z	B	0.5	0.403	0.951	0.427	0.430	0.424
		(0.2)	(0.032)	(0.014)	(0.032)	(0.032)	(0.032)
ρ_Υ	B	0.5	0.192	0.195	0.175	0.177	0.173
		(0.2)	(0.056)	(0.051)	(0.045)	(0.045)	(0.046)
ρ_d	B	0.5	0.955	0.032	0.935	0.927	0.934
		(0.2)	(0.024)	(0.019)	(0.027)	(0.029)	(0.031)
ρ_p	B	0.5	0.959	0.398	0.983	0.985	0.983
		(0.2)	(0.008)	(0.063)	(0.008)	(0.007)	(0.007)
ρ_w	B	0.5	0.992	0.993	0.986	0.985	0.986
		(0.2)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
μ_p	B	0.8	0.785	0.788	0.788	0.787	0.787
		(0.1)	(0.108)	(0.103)	(0.103)	(0.103)	(0.104)
μ_w	B	0.8	0.789	0.784	0.780	0.789	0.786
		(0.1)	(0.101)	(0.107)	(0.114)	(0.102)	(0.105)
σ_g	U	0.1	0.0148	0.0148	0.0148	0.0148	0.0148
		(1.0)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
σ_z	U	0.1	0.0166	0.0035	0.0159	0.0157	0.0157
		(1.0)	(0.0010)	(0.0006)	(0.0009)	(0.0009)	(0.0009)
σ_Υ	U	0.1	0.0473	0.0412	0.0413	0.0415	0.0417
		(1.0)	(0.0032)	(0.0031)	(0.0028)	(0.0028)	(0.0028)
σ_d	U	0.1	0.0344	0.0270	0.0176	0.0169	0.0181
		(1.0)	(0.0135)	(0.0043)	(0.0043)	(0.0039)	(0.0064)
σ_p	U	0.1	0.1170	0.1274	0.1076	0.1081	0.1081
		(1.0)	(0.0225)	(0.0223)	(0.0177)	(0.0178)	(0.0178)
σ_w	U	0.1	1.1460	1.2612	1.3441	1.3488	1.3168
		(1.0)	(0.1942)	(0.2254)	(0.2889)	(0.2198)	(0.2203)
σ_r	U	0.1	0.0025	0.0023	0.0026	0.0026	0.0026
		(1.0)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
σ_τ	U	0.1	0.0260	0.0260	0.0261	0.0261	0.0261
		(1.0)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)

Notes. Table shows prior distributions and Bayesian estimates of parameters across different models. Notation in the second columns is as follows: B = beta, G = gamma, I = inverse gamma distributions. Estimates are presented as mean values and standard deviations across the last $700,000$ out of 1 million elements of a Markov chain generated using the Metropolis Hastings algorithm. Kalman filter is used to evaluate the likelihood of the data.

Table 6: Unconditional second moments in the models and data

	Deep Habits	ROT	G in Utility	Productive G	Baseline	Data
	Standard Deviations					
Government spending	4.79	4.90	4.91	4.76	4.85	7.50
	[3.31, 7.51]	[3.25, 7.92]	[3.38, 7.55]	[3.33, 8.01]	[3.26, 7.97]	
Taxes	7.49	7.26	7.14	7.31	7.24	9.89
	[5.22,11.00]	[5.27,10.62]	[5.20,11.09]	[5.31,10.82]	[5.14,10.73]	
Consumption	2.28	2.07	1.83	1.81	1.85	0.57
	[1.74, 3.61]	[1.73, 2.89]	[1.42, 2.96]	[1.42, 3.03]	[1.41, 3.11]	
Investment	10.18	8.49	9.00	9.13	9.12	3.52
	[7.99,13.54]	[7.03,10.42]	[7.35,11.44]	[7.40,11.49]	[7.39,11.59]	
Hours	15.35	14.83	13.54	13.73	13.66	5.67
	[10.55,22.05]	[9.98,21.99]	[9.30,19.69]	[9.53,19.84]	[9.14,19.88]	
Wages	1.83	1.85	1.53	1.52	1.55	0.59
	[1.47, 2.51]	[1.57, 2.18]	[1.26, 2.21]	[1.24, 2.33]	[1.27, 2.29]	
Inflation	4.38	5.11	3.92	3.90	3.98	2.37
	[3.04, 6.71]	[3.59, 7.21]	[2.89, 5.72]	[2.85, 5.64]	[2.90, 5.70]	
Interest rate	5.37	6.68	4.37	4.37	4.44	3.38
	[3.51, 8.47]	[4.53, 9.71]	[3.04, 6.59]	[3.04, 6.46]	[3.07, 6.62]	
	Standard Deviation Relative to Output Growth					
Government spending	1.83	1.96	2.00	1.94	1.96	7.84
	[1.12,3.02]	[1.23,3.24]	[1.24,3.16]	[1.19,3.28]	[1.21,3.41]	
Taxes	2.84	2.90	2.94	2.95	2.95	10.34
	[1.77,4.55]	[1.98,4.48]	[1.91,4.72]	[1.85,4.65]	[1.91,4.67]	
Consumption	0.87	0.84	0.75	0.74	0.75	0.59
	[0.61,1.35]	[0.66,1.11]	[0.56,1.16]	[0.53,1.18]	[0.55,1.20]	
Investment	3.88	3.42	3.71	3.72	3.72	3.68
	[2.90,4.81]	[2.79,4.03]	[2.84,4.45]	[2.78,4.46]	[2.84,4.51]	
Hours	5.74	5.95	5.47	5.52	5.45	5.93
	[3.81,8.31]	[4.11,8.32]	[3.62,8.07]	[3.73,7.99]	[3.71,8.00]	
Wages	0.70	0.75	0.63	0.62	0.65	0.62
	[0.49,0.95]	[0.58,0.89]	[0.46,0.90]	[0.44,0.91]	[0.47,0.91]	
Inflation	1.62	2.03	1.61	1.57	1.62	2.48
	[1.05,2.59]	[1.44,2.81]	[1.09,2.41]	[1.05,2.37]	[1.09,2.46]	
Interest rate	2.02	2.67	1.79	1.75	1.80	3.53
	[1.25,3.16]	[1.83,3.75]	[1.16,2.75]	[1.11,2.68]	[1.18,2.75]	

Notes: The table shows the standard deviations of the observable variables. The moments reported are the median values of the moment distribution created by generating an artificial sample with the same length as our dataset (225 observations) after discarding the 50 initial observations, for a random sample of 1000 parameter draws from the Markov chain obtained as part of the model estimation procedure. The numbers in the brackets give the 5th and 95th percentile numbers for the moments. Inflation and interest rates are annualized.

Table 7: Unconditional correlations in the models and data

	Deep Habits	ROT	G in Utility	Productive G	Baseline	Data
	Serial correlation					
Government spending	0.95	0.95	0.95	0.95	0.95	0.98
	[0.89,0.98]	[0.90,0.98]	[0.90,0.98]	[0.90,0.98]	[0.89,0.98]	
Taxes	0.93	0.93	0.93	0.93	0.93	0.96
	[0.87,0.97]	[0.87,0.97]	[0.87,0.97]	[0.87,0.97]	[0.86,0.97]	
Consumption	0.53	0.26	0.56	0.55	0.54	0.23
	[0.21,0.73]	[0.10,0.46]	[0.21,0.75]	[0.19,0.77]	[0.19,0.76]	
Investment	0.78	0.74	0.74	0.74	0.74	0.29
	[0.46,0.88]	[0.50,0.84]	[0.48,0.84]	[0.46,0.85]	[0.49,0.84]	
Hours	0.98	0.98	0.98	0.98	0.98	0.97
	[0.97,0.99]	[0.97,0.99]	[0.97,0.99]	[0.97,0.99]	[0.97,0.99]	
Wages	0.47	0.59	0.58	0.58	0.55	0.05
	[0.23,0.66]	[0.44,0.71]	[0.25,0.73]	[0.24,0.73]	[0.23,0.72]	
Inflation	0.93	0.94	0.91	0.91	0.91	0.84
	[0.86,0.97]	[0.88,0.97]	[0.84,0.95]	[0.84,0.95]	[0.84,0.95]	
Interest rate	0.97	0.98	0.96	0.96	0.96	0.95
	[0.94,0.99]	[0.95,0.99]	[0.92,0.98]	[0.92,0.98]	[0.92,0.98]	
	Correlation with Output Growth					
Government spending	0.01	0.02	0.02	0.01	0.02	0.08
	[-0.22,0.24]	[-0.25,0.30]	[-0.20,0.25]	[-0.17,0.23]	[-0.20,0.25]	
Taxes	-0.13	-0.07	-0.10	-0.10	-0.10	-0.09
	[-0.36,0.12]	[-0.33,0.21]	[-0.31,0.12]	[-0.30,0.11]	[-0.30,0.11]	
Consumption	0.23	0.59	0.38	0.36	0.37	0.60
	[-0.15,0.64]	[0.32,0.77]	[0.01,0.70]	[-0.00,0.71]	[-0.01,0.70]	
Investment	0.77	0.73	0.76	0.77	0.76	0.86
	[0.55,0.87]	[0.58,0.83]	[0.59,0.86]	[0.55,0.87]	[0.57,0.86]	
Hours	0.28	0.34	0.24	0.25	0.24	0.08
	[0.14,0.41]	[0.15,0.50]	[0.10,0.37]	[0.11,0.36]	[0.10,0.37]	
Wages	0.51	0.61	0.50	0.48	0.48	0.10
	[0.19,0.73]	[0.34,0.75]	[0.13,0.74]	[0.14,0.74]	[0.09,0.74]	
Inflation	-0.03	0.44	0.08	0.06	0.08	-0.28
	[-0.26,0.21]	[0.28,0.59]	[-0.14,0.30]	[-0.15,0.29]	[-0.15,0.28]	
Interest rate	0.19	0.42	0.19	0.19	0.19	-0.18
	[-0.03,0.41]	[0.26,0.58]	[-0.02,0.40]	[-0.02,0.38]	[-0.03,0.39]	

Notes: The table shows the correlations of the observable variables. The moments reported are the median values of the moment distribution created by generating an artificial sample with the same length as our dataset (225 observations) after discarding the 50 initial observations, for a random sample of 1000 parameter draws from the Markov chain obtained as part of the model estimation procedure. The numbers in the brackets give the 5th and 95th percentile numbers for the moments. Inflation and interest rates are annualized.

Table 8: Contribution of the government spending shock to model volatility, %

	Deep Habits	ROT	G in Utility	Productive G	Baseline
Taxes	0.69	0.84	0.65	0.72	0.70
	[0.22, 2.38]	[0.30, 2.78]	[0.21, 2.18]	[0.22, 2.49]	[0.21, 2.28]
Consumption	6.96	4.63	8.61	7.57	7.74
	[4.22, 10.25]	[2.90, 7.04]	[5.09, 12.88]	[4.44, 11.37]	[4.34, 11.36]
Investment	1.99	2.51	1.04	1.24	1.23
	[1.06, 3.87]	[1.51, 4.18]	[0.30, 1.84]	[0.52, 2.21]	[0.54, 2.05]
Hours	5.74	8.04	6.00	6.07	6.30
	[3.52, 9.92]	[4.88, 14.01]	[3.58, 10.86]	[3.68, 10.86]	[3.74, 11.09]
Wages	2.79	1.91	2.49	2.73	2.66
	[1.35, 3.93]	[0.78, 3.75]	[1.60, 3.39]	[1.73, 3.72]	[1.75, 3.69]
Inflation	1.62	5.61	3.56	3.14	3.03
	[0.69, 10.77]	[3.23, 9.28]	[2.00, 5.84]	[1.69, 5.15]	[1.62, 4.82]
Interest rate	1.96	5.27	3.57	3.30	3.13
	[0.93, 4.52]	[3.05, 8.87]	[1.96, 6.11]	[1.84, 5.43]	[1.80, 5.12]

Notes: The table shows the standard deviations in a model with government spending shock as a ratio of unconditional model implied standard deviation, in percentages. The reported numbers are the median values created by a random subsample of 1000 elements of a Markov chain obtained as part of the model estimation procedure. The numbers in the brackets give the 5th and 95th percentile shares.

Table 9: Model marginal likelihood: DSGE-VARs

p	BL	DH vs. BL	ROT vs. BL	G in U vs. BL	G in F vs. BL
0.1	5383.4	-9.8	-5.0	-7.2	0.9
0.5	5385.5	-9.9	-5.1	-7.3	0.7
0.9	5387.6	-9.8	-5.1	-7.2	0.8

Notes. Table shows logarithm of marginal likelihood of a model evaluated using Geweke (1999) procedure. The first column is the parameter p in the Geweke estimator that specifies the supplementary p.d.f $f(\theta)$ in Equation (17). The second column shows the marginal likelihood in the baseline model. Columns 3-6 present the log of marginal likelihood of a model relative to the best fitted model, so that negative numbers indicate more poor fit. DH = model with deep habits, ROT = for rule-of-thumb model, G in U = model with government spending in the utility function, G in F = the model with government spending in the production technology, and BL = the baseline model with no specific features.

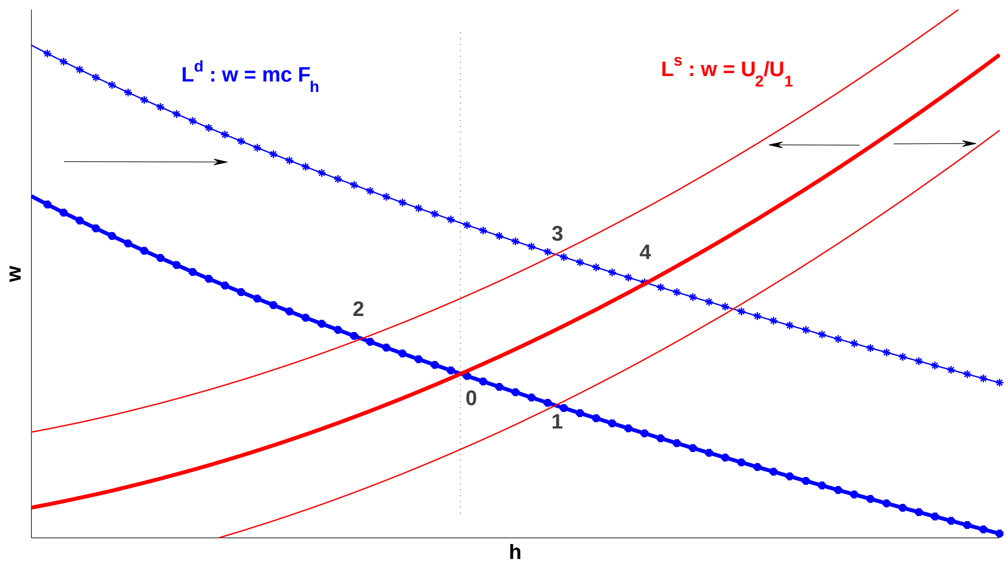

Figure 1: Effect of the government spending shock in labor market.

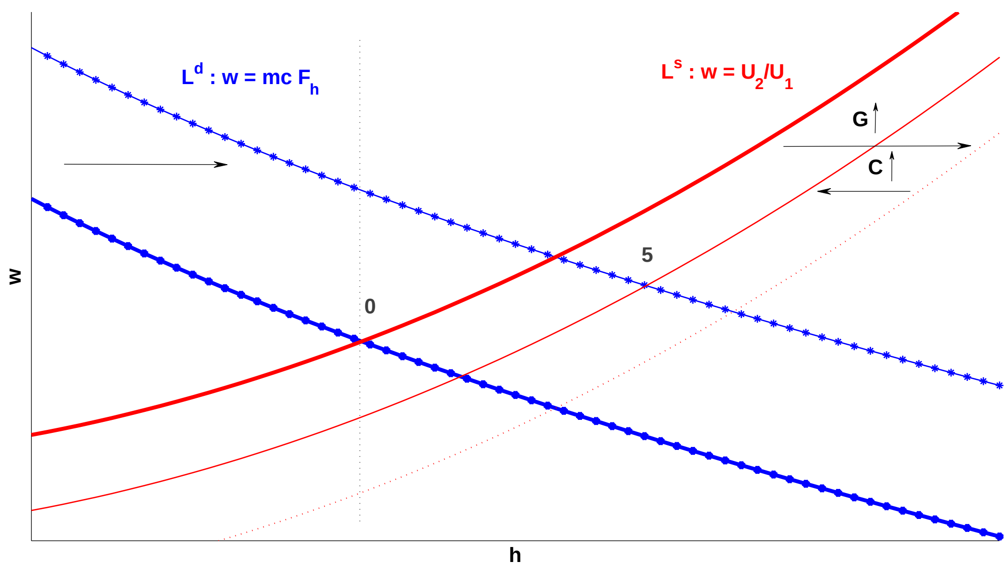

Figure 2: Effect of the government spending shock in labor market in the model where government spending affects utility.

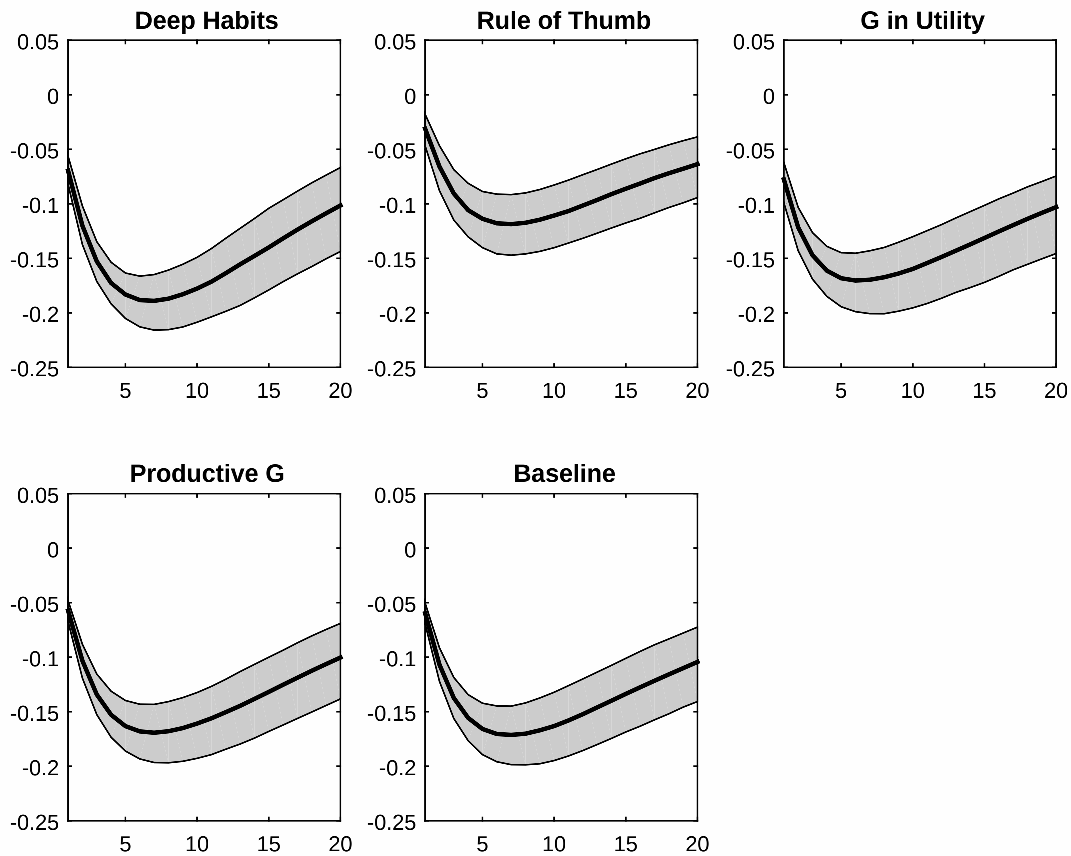

Figure 3: Consumption response to the government spending shock

Notes: Each graph shows an impulse response to a 1 percent government spending shock in percentage deviations from trend. Quarters are along the horizontal axis, and percentages are on the vertical axis. Each response is calculated as the median value of the impulse response distribution created by a random subsample of 1000 elements of a Markov chain obtained as part of the model estimation procedure. The shaded regions show the 5^{th} and 95^{th} quantile of this distribution.

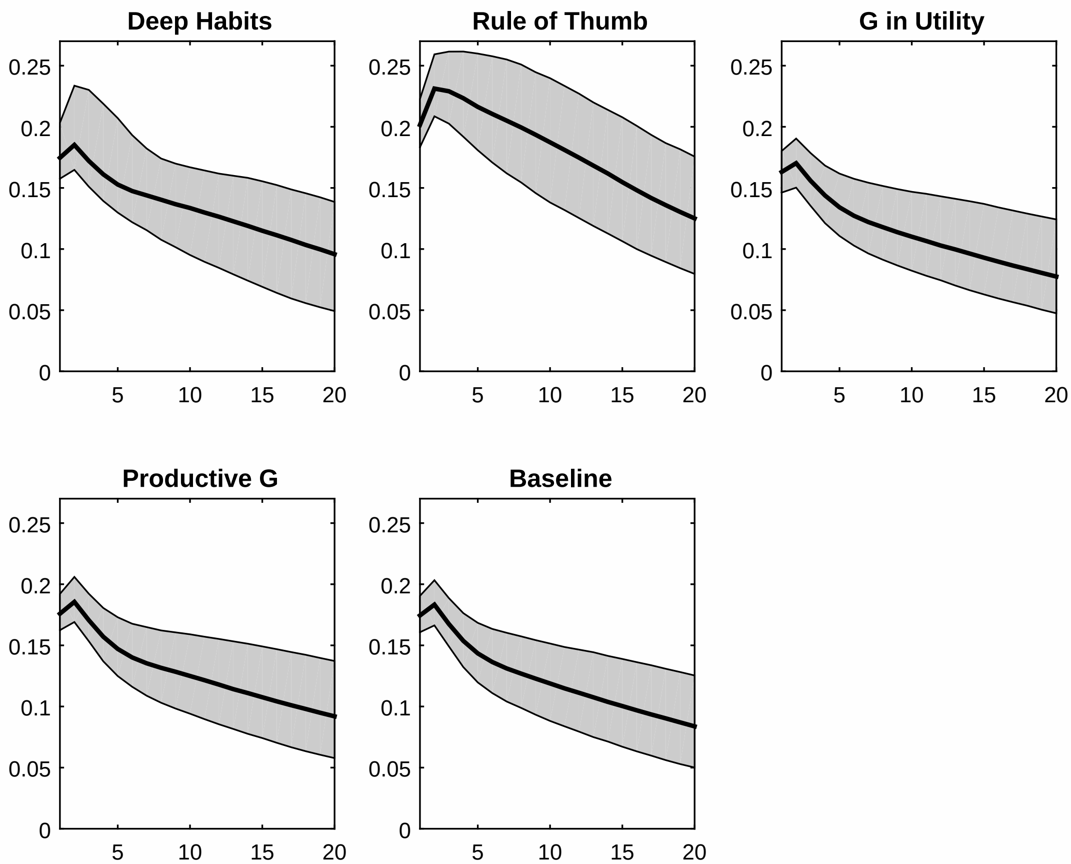

Figure 4: Output response to the government spending shock

Notes: Each graph shows an impulse response to a 1 percent government spending shock in percentage deviations from trend. Quarters are along the horizontal axis, and percentages are on the vertical axis. Each response is calculated as the median value of the impulse response distribution created by a random subsample of 1000 elements of a Markov chain obtained as part of the model estimation procedure. The shaded regions show the 5^{th} and 95^{th} quantile of this distribution.

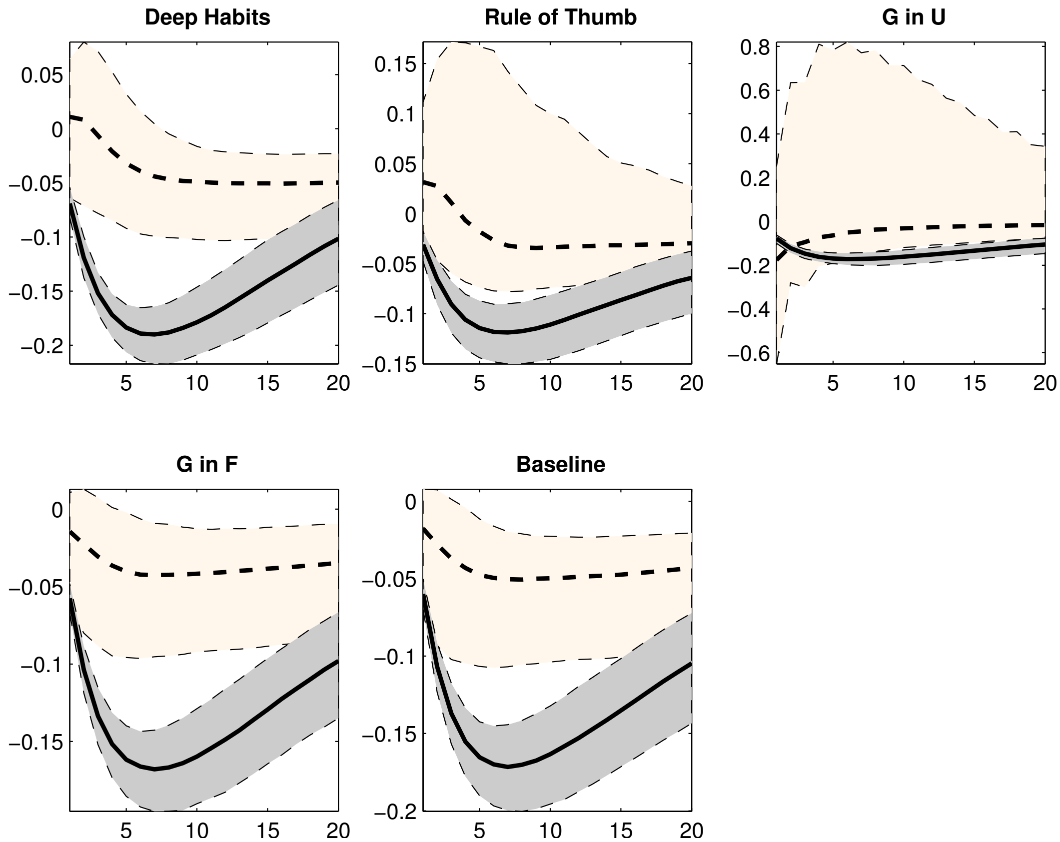

Figure 5: Prior predictive analysis: Consumption response to the government spending shock under priors

Notes: Each graph shows an impulse response to a 1 percent government spending shock in percentage deviations from trend. The dashed lines show the impulse response distribution based on random subsample of 1000 parameter draws from the prior distribution. The solid lines show the impulse response distribution created by a random subsample of 1000 elements of a Markov chain obtained as part of the model estimation procedure. Quarters are along the horizontal axis, and percentages are on the vertical axis. The corresponding shaded regions lines show the 5^{th} and 95^{th} quantile of each distribution.

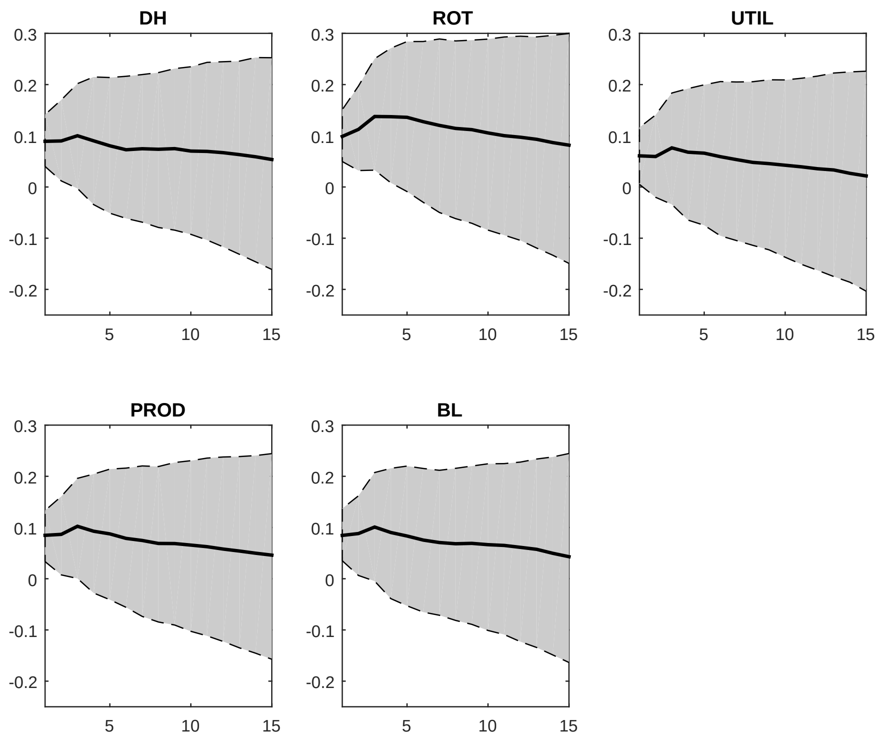

Figure 6: Consumption response to the government spending shock in the DSGE-VAR models

Notes: Each graph shows an impulse response to a 1 percent government spending shock in percentage deviations from trend, based on the estimation of DSGE-VAR model. Quarters are along the horizontal axis, and percentages are on the vertical axis. The corresponding shaded regions show the 5^{th} and 95^{th} quantile of each distribution.