unibz
Fakultät für Wirtschaftswissenschaften
Facoltà di Economia
Faculty of Economics and Management

BEMPS –
Bozen Economics & Management
Paper Series

NO 84/ 2021

Adaptive Importance Sampling for DSGE Models

Stefano Grassi, Marco Lorusso,
Francesco Ravazzolo

Adaptive Importance Sampling for DSGE Models*

Stefano Grassi[†] **Marco Lorusso**[‡] **Francesco Ravazzolo**[§]

May 26, 2021

Abstract

This paper introduces a new adaptive methodology for the estimation of Dynamic Stochastic General Equilibrium (DSGE) models based on the Mixture of Students t by Importance Sampling weighted Expectation-Maximization (MitISEM). The use of Importance Sampling and of an adaptive scheme based on Expectation-Maximization allows us to efficiently estimate any sort of DSGE model. We apply the MitISEM in simulation examples with two workhorse DSGE models. Our results indicate how the MitISEM achieves identification of the model parameters even in the presence of bimodality. We also use the MitISEM to estimate an open economy model encompassing international trade between two countries, namely Canada and the US. For both countries, we consider a rich fiscal policy sector that includes two different types of public expenditure: productive and unproductive government spending. Our findings show that, in the presence of nominal rigidities, an increase in productive spending generates a crowding-in on domestic private consumption, whereas unproductive spending induces a fall in domestic private consumption. We also find that irrespective of the type of government expenditure, an increase in public spending for the domestic economy induces an exchange rate appreciation and an improvement in the trade balance. Finally, our results show that the degree of trade openness matters in terms of propagation of government spending shocks.

Keywords: Adaptive Importance Sampling, DSGE Model, Expectation-Maximization, Fiscal policy, Open-Economy Model.
JEL Classification: C12, C22, G12, G13.

*We would like to thank Herman K. van Dijk and the conference participants at the 14th International Conference on Computational and Financial Econometrics (CFE 2020).

[†]University of Rome Tor Vergata.
[‡]Newcastle University Business School, Department of Economics.
[§]Free University of Bozen-Bolzano; CAMP, BI Norwegian Business School; and RCEA.

1 Introduction

In recent years, Dynamic Stochastic General Equilibrium (DSGE) models have been widely adopted in academia and central banks to study the behaviour of macroeconomic time-series over the business cycle, as well as for policy analysis and forecasting.

There are different types of DSGE models and scholars often disagree on their classification.[1] Initially, DSGE models were real business cycle models that featured dynamics and general equilibrium (Kydland and Prescott, 1982; Long and Plosser, 1983; Campbell, 1994). Successively, new Keyensian DSGE models with nominal frictions were developed (Yun, 1996; Clarida et al., 1999; Woodford, 2003; An and Schorfheide, 2007; Galí, 2015; Herbst and Schorfheide, 2016). Later on, DSGE models were extended to capture the joint dynamics of output, consumption, investment, hours, wages, inflation, and interest rates for a closed economy (Christiano et al., 2005; Smets and Wouters, 2003; Smets and Wouters, 2007).

Without being exhaustive, one strand of literature has analysed the impact of government spending shocks on the economy, with specific attention paid to private consumption. Some DSGE models incorporated rule-of-thumb consumers (Galí et al., 2007; Bilbiie et al., 2008), whereas others included consumption habits (Schmitt-Grohé and Uribe, 2004; Ravn et al., 2006). Some DSGE models assumed that government spending contributes to aggregate production (Baxter and King, 1993; Ambler and Paquet, 1996; Linnemann and Schabert, 2006), whereas others used non-separable utility functions (Linnemann, 2006; Bilbiie, 2009; Bilbiie, 2011; Coenen et al., 2012). There are also DSGE models that have investigated the effects of fiscal policy shocks with the use of fiscal rules (Forni et al., 2009; Leeper et al., 2010b; Asimakopoulos et al., 2020).

Another strand of the literature relates to DSGE models that present open economy features. Some DSGE models have focused on monetary policy in open economies (Corsetti and Pesenti, 2001; Kollmann, 2001; Galí and Monacelli, 2005; Lubik and Schorfheide, 2005). Other DSGE models have adopted a new Keynesian framework of a small open economy (Adolfson et al., 2008; Justiniano and Preston, 2010; Born et al., 2013). There are also multi-country DSGE models that have included nominal and real frictions (Adolfson et al., 2007; Erceg et al., 2008; Rabanal and Tuesta, 2010; Bodenstein et al., 2011), and DSGE models that have linked international macroeconomics and trade theory (Ghironi and Melits, 2005; Cacciatore and Traum, 2020).

Over time, DSGE models have clearly increased their level of complexity. This implies that their

[1]For example, Ghironi (2017) and Blanchard (2017) propose different criteria in order to classify DSGE models.

estimation has become a more challenging task. Starting with Schorfheide (2000) and Otrok (2001), Markov Chain Monte Carlo (MCMC) algorithms, and more specifically the random walk Metropolis-Hastings (RWMH), have been the cornerstones for DSGE estimation. The easy implementation of the RWMH, together with the *ad hoc* program Dynare (see Adjemian et al., 2011), has contributed to the massive spread of this estimation approach. Herbst (2012) reported that 95% of papers published from 2005 to 2010 in the eight top economics journals use the RWMH algorithm for DSGE estimation. However, as DSGE complexity increased, the limit of the RWMH estimation approach emerged. Chib and Ramamurthy (2010) and Herbst (2012) have documented that the RWMH is very slow to converge and can get stuck near the local mode. Moreover, it can be very autocorrelated, resulting in an inefficient estimation, and fails to explore the entire posterior distribution (see, for example, Herbst and Schorfheide, 2014). These shortcomings have called for different estimation approaches.

Starting from the results in Creal (2007), recently Herbst and Schorfheide (2014) and Cai et al. (2020) proposed the application of the Sequential Monte Carlo (SMC) algorithm for DSGE estimation. The SMC is usually associated with filtering non-linear-state-space systems and it was originally proposed by Herbst and Schorfheide (2014). It recursively constructs importance samplers for a sequence of distributions that begin at an easy-to-sample initial distribution and that end at the posterior. The SMC is a flexible methodology that has also been applied in the Markow-Switching VAR, see Bognanni and Herbst (2018).

In this paper we take a different approach and, in particular, we propose a new methodology that is based on the Mixture of (Student's) t by Importance Sampling weighted Expectation-Maximization (MitISEM). Our work presents two main contributions with respect to the previous literature. Firstly, we propose this new estimation strategy that can easily tackle complex DSGE models. Secondly, we provide novel insights into the role of fiscal policy in an open economy DSGE model. In particular, we apply the MitISEM methodology to a two-country DSGE model that aims to analyse the effects of government spending shocks on the economy. Our DSGE model presents new Keynesian features and consider two types of government expenditures, namely, productive and unproductive spending.

The MitISEM methodology can be summarised as follows. After an initial set of candidate draws, the algorithm applies Important Sampling (IS) to compute the unknown posterior density by a mixture of Student's t densities. Importance weights emphasize certain values of the posterior distributions, which achieve identification of the posterior densities. Moreover, at each simulation, the parameters of the mixture of Student's t densities adapt to the most recent IS draws and are recomputed via an Expectation-Maximization (EM) step.

Our adaptive scheme provides large benefits. Firstly, on the computational side, the algorithm is "embarrassingly parallelizable" on multiple processors or graphics processing units (GPU) (see, for example, Baştürk et al., 2016). This allows any user to estimate complex DSGE models in a reasonable computing time. Secondly, the algorithm does not require the parameters to be tuned for the user. This is particularly important when the number of parameters is very large, and their distribution is non-standard. Thirdly, the adaptive EM step improves and speeds up posterior convergence in the case of parameter identification. Therefore, the posterior estimation offers a tool for investigating whether the parameters can be identified or not. Fourthly, the algorithm relies on the mixture of the Student's t densities, that can easily handle the asymmetry, non-normality and multi-modality of the posteriors. This aspect is particularly important when the likelihood function is complex and not well distributed. Finally, the MitISEM extends the RWMH method by developing a new algorithm that does not rely on any Metropolis-Hastings (MH)step. This reduces the correlation among draws, improves the convergence, and the accuracy of the posterior estimates.

Simulation results show how the MitISEM achieves identification of the model parameters and how it can estimate complex features, such as parameter bimodality. Then, we use the MitISEM to estimate two workhorse DSGE models: the small new Keynesian (NK) model in the spirit of Woodford (2003) and the Smets and Wouters (2007) (SW) model. We compare the estimates obtained using the MitISEM with those of the standard MCMC. We find that the differences in the values of estimated parameters are negligible. In addition, MitISEM presents an enormous advantage in terms of computing time.

As a next step, we apply the MitISEM to a more complex DSGE model. In particular, we estimate a new Keynesian DSGE model that has a two-country framework and is based on 164 equations with 86 parameters. As explained above, the MitISEM presents numerous advantages compared to the RWMH in the estimation of this type of complex DSGE model.

From a theoretical point of view, our DSGE model extends the work by Erceg et al. (2008) by assuming two different types of public expenditure, namely, productive and unproductive government spending. In line with Leeper et al. (2010a) and Asimakopoulos et al. (2020), we include distortive taxes to capital and labour incomes, together with several fiscal policy rules. Our analysis provides insights into two topics that have been widely studied by the previous literature. The first topic relates to the effects of government spending shocks on private consumption. Earlier studies have found contrasting results regarding such effects (see, for example, Blanchard and Perotti, 2002; Fatás and Mihov, 2003; Mountford and Uhlig, 2009; Perotti, 2014; Galí et al., 2007). We find that, in the presence of nominal rigidities, an increase in productive spending generates a crowding-in effect on domestic private consumption. On the other hand, an increase in unproductive government spending induces a crowding-out effect on domestic private consumption even in the presence of nominal rigidities.

Our empirical results also shed light on the effects of government spending shocks on the real exchange rate. Previous papers on this topic have found mixed results (Corsetti and Müller, 2006; Kim and Roubini, 2008; Enders et al., 2011; Bouakez et al., 2014; Beetsma et al., 2008; Born et al., 2013; Ilzetzki et al., 2013; Auerbach and Gorodnichenko, 2016). We show that irrespective of the type of government expenditure, an increase in public spending for the domestic economy induces an exchange rate appreciation and an improvement in the trade balance.

In addition, we find that output multipliers for the domestic economy are larger in the presence of nominal rigidities than when prices and wages are fully flexible. Moreover, under the economy with nominal rigidities, we observe positive consumption present-value multipliers in response to a productive government spending shock. We also show that government spending shocks have different effects on output and consumption multipliers depending on the degree of trade openness of the economy. More specifically, in the presence of nominal rigidities, we observe higher output and consumption multipliers when the elasticity of substitution between domestic and foreign goods is larger. In the case of the economy with flexible prices and wages we find the opposite result.

This paper is organized as follows. Section 2 introduces the MitISEM and shows how this methodology deals with the parameter identification and complex parameter estimation, such as bimodality. In this section we also apply the MitISEM to the NK and the SW models. Section 3 describes the open-economy DSGE model and shows the estimation results obtained using the MitISEM. Finally, Section 4 draws some conclusions.

2 MitISEM for DSGE estimation

This section introduces a new estimation approach for DSGE models based on the tailored construction of the proposal distribution within an IS algorithm. In particular, we adapt the MitISEM introduced by Hoogerheide et al. (2012) and improved in Baştürk et al. (2017) to estimate large DSGE models.[2] Firstly, we describe the algorithm. Secondly, we present how the method works with relatively simple exercises in terms of parameter identification and estimation. Thirdly, we carry out simulation examples with two workhorse DSGE models to demonstrate the accuracy of the MitISEM method. Finally, we propose a robustness experiment.

2.1 MitISEM estimation algorithm

The algorithm is given by the following steps in order to obtain an approximation of a target density, i.e., the unknown parameter posteriors. Let's assume that the objective is to estimate a vector of parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(N)})$ from the unknown posterior $p(\boldsymbol{\theta}|Y)$.

1. **Initialization**: Simulate parameters draws $\boldsymbol{\theta}_0 = (\boldsymbol{\theta}_0^{(1)}, \dots, \boldsymbol{\theta}_0^{(N)})$ from a 'naive' Student-t candidate distribution, g_{naive}:

$$g_{naive} \sim t(\mu_0, \Sigma_0, \nu_0) \tag{1}$$

   where μ_0 and Σ_0 are the mean and scale matrix of the Student-t distribution, respectively. They are computed from a preliminary maximization of the log kernel posterior density (equal to log-priors plus log-likelihood) evaluated at the mode. Therefore, the initialization depends on both the prior assumption and the likelihood. The degrees of freedom ν_0 are a-priori chosen from the user. We suggest applying a low value in order to allow for fat tails, for example $\nu_0 = 3$. Moreover, we apply the same degrees of freedom for all parameters, but this assumption can be relaxed. Furthermore, in the robustness experiment of Section 2.4, we propose an alternative hierarchical approach where μ_0 and Σ_0 are not fixed to their optimal values, but they are drawn from prior distributions. Results are qualitatively similar in the two cases.

   All the parameters are drawn jointly from g_{naive} and simulations are independent across draws. This step corresponds to an independent Metropolis-Hastings step where the candidate is the g_{naive} distribution and the acceptance rate is 1.

[2]See the R library in Baştürk et al. (2017) for applications to financial data.

2. **Adaptation**: Estimate the target distribution's mean (μ^0_{Adap}) and covariance matrix (Σ^0_{Adap}) by applying an IS method to the draws $\boldsymbol{\theta}_0^{(1)}, \ldots, \boldsymbol{\theta}_0^{(N)}$ from g_{naive} in step 1.[3] IS emphasizes certain areas of the importance distribution g_{naive} by sampling more frequently from these values. The importance weights, computed as the ratio between the target distribution and the importance distribution

$$\omega_{\boldsymbol{\theta}_0} = \frac{p(\boldsymbol{\theta}_0)|Y}{g_{naive}(\boldsymbol{\theta}_0|Y)} \tag{2}$$

are used to generate a new sample of draws $\boldsymbol{\theta}_{0,Adap}^{(1)}, \ldots, \boldsymbol{\theta}_{0,Adap}^{(N)}$ from:

$$g^0_{Adap} \sim t(\mu^0_{Adap}, \Sigma^0_{Adap}, \nu). \tag{3}$$

Compute the IS weights $\omega_{\boldsymbol{\theta}_{Adap}}$ for this sample. The basic methodology in IS is to choose a distribution which "encourages" the important values. If this happens, the target (posterior distribution of the parameters) will have smaller variance, and the parameters will be more precisely identified.

Moreover, at each simulation, the parameters of the mixture of Student's t densities adapt to the most recent IS draws and are recomputed via an Expectation-Maximization (EM) step.

3. **IS-weighted EM algorithm**: Apply the Expectation-Maximization (EM) algorithm (see Appendix) given the latest IS weights and draws from step 2. The previous draws are used to derive the new candidate density g^h_{Adap} which is a mixture of Student-t densities:

$$g^h_{Adap} = \sum_{h=1}^{H} \eta_h t_h(\mu_h, \Sigma_h, \nu_h) \tag{4}$$

with optimized mean (μ_h), covariance (Σ_h), degrees of freedom (ν_h) and mixture weight (η_h) computed using an EM algorithm on IS weights and draws from step 2. In the first Monte Carlo draw $H = 1$ (there is only one component) and $\eta_h = 1$ (the only component takes all the weights).

Draw a new sample $\boldsymbol{\theta}_{h,Adap}^{(1)}, \ldots, \boldsymbol{\theta}_{h,Adap}^{(N)}$ from the distribution that corresponds with this proposal density and compute corresponding IS weights.

4. **Iterate on the number of mixture components**: Given the current mixture of H components with corresponding μ_h, Σ_h, ν_h and η_h, $h = 1, \ldots, H$, take a percentage (%) of the

[3] For more details, see Section 3.3. in Robert and Casella (1999).

sample $\boldsymbol{\theta}_{h,Adap}^{(1)}, \ldots, \boldsymbol{\theta}_{h,Adap}^{(N)}$ that corresponds to the highest IS weights. Construct a new mode μ_{H+1} and scale matrix Σ_{H+1} with these draws and IS weights, which are the starting values for the additional component in the mixture candidate density in equation (4). This step ensures that the new component covers a region of the parameter space in which the previous candidate mixture had a relatively low probability mass. Usually, two or three components are sufficient given the flexibility of the mixture of Student's t densities. In the case that the maximum number of components chosen a priori is reached or the convergence in step 5 is achieved, the drawing of a new component is skipped.

Given the latest IS weights and the draws from the current mixture of H components, apply the EM algorithm to optimize (again) each mixture component μ_h, Σ_h, ν_h and η_h with $h = 1, \ldots, H+1$. Draw a new sample from the mixture of $H+1$ components and compute the corresponding IS weights.

5. **Assess convergence of the candidate density's quality by inspecting the IS weights** and return to step 3 unless the algorithm has converged.

One of the advantages of the MitISEM is the adaptation step 2 , which eliminates the extreme dependence of results on user-specified values, especially in the case that the user-specified values are not accurate, and results in higher levels of robustness.[4] Moreover, the algorithm is not really dependent on the initial mode and scale of the parameters (μ_0, Σ_0). These are usually obtained through grid-search algorithms that can incur in local maxima and not positive-definite Hessians. In these cases, the user can specify a reasonable starting μ_0 and Σ_0 that will be updated in the adaptation step or apply a hierarchical prior approach (see Section 2.4).

Step 2 can be seen as an intermediate step that quickly attempts to improve the initial candidate density g_{naive}. If during the EM algorithm, a scale matrix Σ_h of a Student-t component becomes (nearly) singular, then this h-th component is removed from the mixture. Moreover, if during the EM algorithm, a weight η_h becomes very small, then this h-th component is removed from the mixture.

Convergence in step 4 can be assessed by computing the relative change in the Coefficient of Variation (CoV) of the IS weights, i.e., the standard deviation of the IS weights divided by their mean, as in Hoogerheide et al. (2012). The default convergence in MitISEM is defined as the change of the CoV being smaller than 10%, but the user can specify convergence in terms of the

[4]For more details see Hoogerheide et al. (2012) and Baştürk et al. (2017).

acceptance probability. The convergence tolerance can also be changed by the user.

Finally, the specified starting values for ν_{H+1} and η_{H+1} in step 4 are fixed in our exercises to 1 and 0.10, i.e., the new component has fat tails, and a relatively low probability ex-ante.

2.2 Identification experiments

This subsection proposes two identification experiments. Experiment 1 considers the same model as in Andrews and Mikusheva (2015), in which the first set of parameters is identified, whereas the second is not. This exercise explains how MitISEM provides an accurate and precise posterior when the parameters are identified, but not when they are not identified. Experiment 2 shows that the MitISEM is able to capture multi-modality in a stylized state-space model.

Experiment 1. The series y_t is simulated from:

$$y_t = (\pi + \beta)y_{t-1} + \varepsilon_t - \pi\varepsilon_{t-1}, \qquad \varepsilon_t \sim \mathcal{N}(0, 1) \tag{5}$$

where the parameters (β, π) satisfy the following restrictions: $\beta \neq 0$, $|\pi| < 1$, and $|\pi + \beta| < 1$, guaranteeing that the process is stationary and invertible. The model can be rewritten as:

$$(1 - (\pi + \beta)\mathrm{L})y_t = (1 - \pi\mathrm{L})\varepsilon_t.$$

However, the parameter π is not identified when $\beta = 0$. We simulate data in two cases: the first case sets the value of $\beta = 0$, that is π is not identified; the second case fixes $\beta = 0.5$, with the parameter π identified by construction.

Figure 1 shows the estimated posterior distributions of the parameters π and β under-identification ($\beta \neq 0$) and non-identification ($\beta = 0$), for two sample sizes, $T = 250$ and $T = 1000$, respectively. We centred the posterior densities of both β and π at zero in order to make them comparable. From Figure 1 we observe that, as the sample size rises, there is an increase in the precision of the estimation for the identified case. As we discussed above, in step 4 of the algorithm, the posterior densities become very narrow around the true values and precision is high with few simulations. This result is not true for the non-identified case where the posterior densities are very diffuse. Therefore, the MitISEM can also offer a useful tool for investigating whether a parameter is identified or not.

Experiment 2. The MitISEM is based on a mixture of Student's t, therefore, it can capture bimodality in the parameter space. To test this point we replicate the same exercise as in Herbst and Schorfheide (2014), where the posterior density $p(\theta|Y)$ of a stylized state-space model is calculated. The model is given by:

$$y_t = \begin{bmatrix} 1 & 1 \end{bmatrix} s_t, \quad s_t = \begin{bmatrix} \phi_1 & 0 \\ \phi_3 & \phi_2 \end{bmatrix} s_{t-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1). \tag{6}$$

The mapping between some structural parameters $\mathbf{\Theta} = [\theta_1, \theta_2]$ and the reduced form parameters $\mathbf{\Phi} = [\phi_1, \phi_2, \phi_3]$ is assumed to be:

$$\phi_1 = \theta_1^2, \quad \phi_2 = (1 - \theta_1^2), \quad \phi_3 - \phi_2 = -\theta_1\theta_2.$$

This representation shows two main problems. The first problem is the lack of identifiability of θ_2 when θ_1 is close to 0. This is because these parameters enter the model multiplicatively. The second problem relates to a global identification issue. The root cancellation in the AR and MA lag polynomials for y_t causes a bimodality in the likelihood function.

We simulate $T = 200$ observations from the model given in Equation (6) with the structural parameters set at $\mathbf{\Theta} = [0.45, 0.45]$. Those values are observationally equivalent to $\mathbf{\Theta} = [0.89, 0.22]$. The MitISEM requires that the starting values are specified for each new component, i.e., $\nu_h = 1$ and $\eta_h = 0.10$, $h = 1, \ldots, H$ and $H = 3$. Finally, we use uniform priors, i.e., $0 \leq \theta_1 \leq 1$ and $0 \leq \theta_2 \leq 1$.

The MitISEM algorithm works very well for this simple problem and it is computationally very fast with 10,000 draws executed in a few seconds. Figure 2 shows one of the crucial features of the algorithm: the bimodality is well captured using a mixture of two Student's t densities.

2.3 Simulation examples

This section presents a Monte Carlo experiment that compares the MitISEM to the standard MCMC method. In particular, we compare the estimates of two workhorse DSGE models obtained with the MitISEM and MCMC approaches, respectively.

The first model is the small new Keynesian (NK) model, as described in Woodford (2003), An and Schorfheide (2007), Galí (2015) and Herbst and Schorfheide (2016). This model consists of a representative household, a firm producing final goods, a continuum of firms producing intermediate goods, a central bank and a fiscal authority. We refer to Herbst and Schorfheide (2016) for a detailed description of the model. The set of parameters is given by:

$$\boldsymbol{\Theta} = \{\tau, \kappa, \psi_1, \psi_2, r^{(A)}, \pi^{(A)}, \gamma^{(Q)}, \rho_r, \rho_g, \rho_z, \sigma_r, \sigma_g, \sigma_z\}$$

where individual parameters are defined in Table 1. In the Monte Carlo exercise, we simulate $I = 100$ times the series. Then, for each simulation we estimate the NK model using MCMC and MitISEM and compare the parameter estimation bias. In order to obtain a fair comparison between the two approaches, we use the same number of draws, $D = 200000$. Table 1 reports the bias for the MCMC (B-MCMC) and the bias for the MitISEM (B-MitISEM). Our results indicate that the MitISEM performs well and is very precise in the parameter estimation.

The second model is the popular Smets and Wouters (2007) (SW) model. The theoretical framework consists of a representative household that has a utility function that includes two arguments, consumption and labour. This representative household has monopoly power over wages. This implies the presence of sticky nominal wages à la Calvo (1983). The representative household rents capital services to intermediate production firms and decides how much capital to accumulate given certain capital adjustment costs. As the rental price of capital changes, the utilization of the capital stock can be adjusted at increasing cost. There is a single final production good and a continuum of intermediate production goods. The intermediate production firms produce under monopolistic competition and use capital services and labour as input factors. These intermediate production firms set prices according to the Calvo (1983) model. As additional assumption concerning nominal rigidities, there is partial indexation of both wages and prices to past inflation rates. Given the fact that this model has been widely analysed by previous studies, we omit the presentation of the several equations. The set of parameters is given by:

$$\boldsymbol{\Theta} = \{\varphi, \sigma_c, h, \xi_w, \sigma_l, \xi_p, \iota_w, \iota_p, \psi, \Phi, r_\pi, \rho, r_y, r_{\Delta_y}, \bar{\pi} 100\left(\beta^{-1} - 1\right),$$
$$\bar{l}, \bar{\gamma}, \alpha, \rho_a, \rho_b, \rho_g, \rho_i, \rho_r, \rho_p, \rho_w, \mu_p, \mu_w, \sigma_a, \sigma_b, \sigma_g, \sigma_i, \sigma_r, \sigma_p, \sigma_w\}$$

where individual parameters are defined in Table 2. As above, we simulate the series $I = 100$ times and at each iteration we estimate the model using MitISEM and MCMC. In order to have a fair comparison between the two approaches, we use the same number of draws, $D = 200000$.Table 2 shows that the MitISEM provides accurate estimates of the parameters.[5]

[5]Estimated results for the parameters of the SW model are available from the authors upon request.

Finally, Table 3 provides evidence of the large computational gains of the MitISEM with respect to the MCMC algorithm. In particular, the ratios in the computing time of the MitISEM with respect to the MCMC indicate gains up to 3-4 time when using a large set of cores. Table 3 also shows that when the number of components increases the computational time increases and gains are larger for more complex models.

2.4 Robustness experiment

As we discussed in Section 2.1, the MitISEM requires that both μ_0 and Σ_0 be specified in step 1. As explained in Hoogerheide et al. (2012), a valuable strategy is to maximize the log-posterior density and to use the estimated parameters for μ_0 and minus the inverse of the Hessian matrix for Σ_0. This approach has been proved to be very reliable in financial problems and in microeconometrics with non-elliptical density contours.[6] In the DSGE optimization, it is not rare to incur multimodality and local maxima. In these cases, the estimated μ_0 and Σ_0 may not be the best choice. Moreover, one might not have access to or believe in the initial values of μ_0 and Σ_0. Therefore, a hierarchical approach can be used in order to form a prior knowledge for the two parameters.

To test the robustness of the MitISEM to the initial conditions, we estimate $I = 100$ times the SW model with two initializations. The first one maximizes the log-posterior and fixes μ_0 to the estimated parameters and Σ_0 to minus the inverse of the Hessian matrix; the second simulates from the standard prior of SW and uses their mean as μ_0, and the variance-covariance matrix as Σ_0. The second initialization is similar to the one used in Herbst and Schorfheide (2014).

We compute the mean difference ($M.D.$) and standard deviation difference ($S.D.D.$) across runs defined as:

$$M.D. = \frac{1}{I}\sum_{i=1}^{I}\left(\hat{\mathbf{\Theta}}_i^{LP} - \hat{\mathbf{\Theta}}_i^{Prior}\right), \qquad S.D.D. = \frac{1}{I}\sum_{i=1}^{I}\left(std(\mathbf{\Theta})_i^{PL} - std(\mathbf{\Theta})_i^{Prior}\right). \tag{7}$$

where $\mathbf{\Theta}_i^{LP}$ are the estimated parameters at the i-th iteration of the MitISEM with log-posterior density maximization, and $\mathbf{\Theta}_i^{Prior}$ are the estimated parameter at the i-th iteration of the MitISEM with a hierarchical prior structure.

Table 4 indicates that the differences are very small and, more importantly, the S.D.D. is below the value of 0.1 in almost all the cases when prior values are in single units.

[6] For more details, see Baştürk et al. (2017).

3 An open-economy model

In this section, we develop a more complex DSGE model and we estimate it using the MitISEM. In particular, we aim to analyse the effects of government spending shocks in an open-economy model. Our theoretical framework encompasses international trade between two countries, namely Canada and the US. We assume that these two countries differ in size but are otherwise symmetric. In each country, the representative household maximises its utility function that has two arguments, consumption and labour. The representative household makes investment decisions, owns the capital stock and rents it to intermediate production firms. Each country produces a single final production good and a continuum of intermediate production goods. As in Asimakopoulos et al. (2020), the intermediate production firms produce under monopolistic competition and use three input factors, i.e., private capital, government capital and labour. Nominal rigidities in each country consist of sticky prices and wages à la Calvo (1983), as well as partial indexation of both wages and prices to past inflation rates. In each country, final consumption goods, as well as investment goods, are produced by firms that combine domestic and imported goods under perfect competition. As in Bodenstein et al. (2011), we assume that asset markets are complete at the country level, but are incomplete internationally. In both countries we assume a rich fiscal sector that includes different types of government expenditures, namely productive and unproductive government spending. Moreover, we consider several fiscal policy rules.

Since the model is symmetric, in what follows we will describe only the model for the domestic country in detail , which is Canada in our study.

3.1 Theoretical Model

Households. In each country, the representative household maximizes its lifetime utility function by choosing purchases of consumption, $C_{1,t}$, and investment goods, $I_{1,t}$, capital stock, $K_{1,t}$, and next period's holdings of both domestic government bonds, $B_{1,t+1}$, and foreign government bonds, $B^f_{1,t+1}$, given its period-by-period budget constraint. Therefore, the representative household maximizes:

$$\max_{\left\{C_{1,t},I_{1,t},K_{1,t},B_{1,t},B^f_{1,t}\right\}} \mathbb{E}_t \left\{ \sum_{t=0}^{\infty} \beta_1^t \left[\frac{1}{1-\sigma_1^c} \left(C_{1,t} - h_1 C_{1,t-1}\right)^{1-\sigma_1^c} \exp\left(\frac{\sigma_1^c - 1}{1+\sigma_1^l} \left(L_{1,t}\right)^{1+\sigma_1^l} \right) \right] \right\}, \quad (8)$$

subject to the budget constraint:

$$P^c_{1,t}C_{1,t} + P^i_{1,t}I_{1,t} + \left(R^b_{1,t}\right)^{-1} B_{1,t+1} + \frac{e_{1,t}\left(R^b_{2,t}\right)^{-1} B^f_{1,t+1}}{\phi^b_{1,t}} \tag{9}$$
$$= B_{1,t} + e_{1,t}B^f_{1,t} + (1-\tau^l_{1,t})W_{1,t}L_{1,t} + (1-\tau^k_{1,t})R^k_{1,t}K_{1,t-1} + D_{1,t} + T_{1,t},$$

and the capital accumulation equation:

$$K_{1,t} = (1-\delta_1)\,K_{1,t-1} + \varepsilon^i_{1,t}\left(1 - S\left(\frac{I_{1,t}}{I_{1,t-1}}\right)^2\right) I_{1,t}. \tag{10}$$

In equation (8), $\mathbb{E}_t$ denotes the expectation operator at time t and β^t_1 is the discount factor. The representative household consumption is influenced by the presence of external habit, h_1, related to aggregate past consumption. The parameter σ^c_1 is the coefficient of relative risk aversion. The variable $L_{1,t}$ represents hours worked, while σ^l_1 is the inverse of the elasticity of work with respect to the real wage.

In equation (9), $P^c_{1,t}$ and $P^i_{1,t}$ indicate the prices of consumption and investment goods, respectively. The gross nominal return of the domestic government bond is denoted by $R^b_{1,t}$, while $R^b_{2,t}$ is the gross nominal return of the foreign government bond. The latter is denominated in foreign currency and, thus, its domestic value depends on the nominal exchange rate, $e_{1,t}$, expressed in units of the domestic currency per unit of foreign currency. As in the paper of Erceg et al. (2008), we assume that the representative household faces an intermediation cost to purchase the foreign bond, $\phi^b_{1,t}$. We indicate by $W_{1,t}$ the aggregate nominal wage, while $R^k_{1,t}$ is the rental rate for capital services. $D_{1,t}$ represents the dividends paid by production goods firms that are owned by the representative household. Moreover, the fiscal authority absorbs part of the gross income of the representative household in order to finance its expenditure. Accordingly, in equation (9), $\tau^l_{1,t}$ denotes the labour income tax rate, while $\tau^k_{1,t}$ is the capital income tax rate. Moreover, $T_{1,t}$ indicates the lump-sum transfers from the government.

The capital accumulation equation (10) includes the adjustment cost function $S(\cdot)$ and an investment specific technology shock denoted by $\varepsilon^i_{1,t}$. Finally, δ_1 denotes the depreciation rate.

We also assume that the representative household has monopoly power over wages that implies sticky nominal wages à la Calvo (1983). Finally, we allow for a partial indexation of wages to past inflation rates.

Firms: production of consumption goods. In each country, the final consumption good, $C_{1,t}$, is produced under perfect competition and sold to the representative household. The representative firm producing final consumption goods uses a constant elasticity of substitution production function. In particular, domestic, $C^d_{1,t}$, and foreign, $M^c_{1,t}$, intermediate consumption goods are combined in order to obtain final consumption goods. The cost minimization problem faced by the representative firm producing final consumption goods is given by:

$$\min_{\{C^d_{1,t}, M^c_{1,t}\}} P^d_{1,t} C^d_{1,t} + P^m_{1,t} M^c_{1,t}$$
$$s.t. : C_{1,t} = \left((\omega^c_1)^{\frac{\rho^c_1}{1+\rho^c_1}} \left(C^d_{1,t} \right)^{\frac{1}{1+\rho^c_1}} + (\omega^{mc}_1)^{\frac{\rho^c_1}{1+\rho^c_1}} \left(\varepsilon^m_{1,t} M^c_{1,t} \right)^{\frac{1}{1+\rho^c_1}} \right)^{1+\rho^c_1} . \tag{11}$$

We denote by ω^c_1 and ω^{mc}_1 the weights of domestic and foreign consumption goods. Moreover, ρ^c_1 represents the elasticity of substitution between domestic and foreign intermediate goods. We assume that import preferences are driven by an exogenous shock, $\varepsilon^m_{1,t}$, that has the form of an AR(1) process. The Lagrange multiplier associated with the cost minimization problem of the representative firm producing final consumption goods is defined as the price of consumption goods $P^c_{1,t}$.

Firms: production of investment goods. Firms producing investment goods, $I_{1,t}$, use a nested constant elasticity of substitution production function. These firms operate under perfect competition and sell investment goods to the representative household. In particular, domestic and foreign investment goods, denoted respectively by $I^d_{1,t}$ and $M^i_{1,t}$, are combined in order to obtain final investment goods. We can express the cost minimization problem of typical firm producing investment goods as follows:

$$\min_{\{I^d_{1,t}, M^i_{1,t}\}} P^d_{1,t} I^d_{1,t} + P^m_{1,t} M^i_{1,t}$$
$$s.t : I_{1,t} = \left(\left(\omega^i_1\right)^{\frac{\rho^i_1}{1+\rho^i_1}} \left(I^d_{1,t} \right)^{\frac{1}{1+\rho^i_1}} + \left(\omega^{mi}_1\right)^{\frac{\rho^i_1}{1+\rho^i_1}} \left(\varepsilon^m_{1,t} M^i_{1,t} \right)^{\frac{1}{1+\rho^i_1}} \right)^{1+\rho^i_1} , \tag{12}$$

where ω^i_1 and ω^{mi}_1 indicate the weights of domestic and foreign investment goods. The elasticity of substitution between domestic and foreign goods is denoted by ρ^i_1. We also assume that investment goods are influenced by an import preferences shock, $\varepsilon^m_{1,t}$, that is the same we assumed in the production of consumption goods. The Lagrange multiplier associated with the problem of cost minimization of the typical investment goods firm coincides with the price of investment goods $P^i_{1,t}$.

Firms: production of domestic intermediate goods. Each country produces a single final production good and a continuum of intermediate production goods. Each intermediate good firm j produces its differentiated output using the Cobb-Douglas technology with three input factors, i.e., private capital ($K_{1,t}$), labour ($L_{1,t}$) and productive government capital ($K^{gp}_{1,t}$):

$$\begin{aligned} &\min_{\left\{K_{1,t}(j),K^{gp}_{1,t}(j),L_{1,t}(j)\right\}} \left(R^k_{1,t}K_{1,t}\left(j\right)+W_{1,t}L_{1,t}\left(j\right)+P^{kg}_{1,t}K^{gp}_{1,t}\left(j\right)\right) \\ &s.t.:Y_{1,t}\left(j\right)=\varepsilon^a_{1,t}\left(K_{1,t}\left(j\right)\right)^{\alpha^k_1}\left(L_{1,t}\left(j\right)\right)^{\alpha^l_1}\left(K^{gp}_{1,t}\left(j\right)\right)^{\alpha^{kg}_1} \\ &\text{where}:\ \alpha^k_1+\alpha^l_1=1 \\ &\text{and}:\ 0<\alpha^{kg}_1<1, \end{aligned} \tag{13}$$

where α^k_1 and α^l_1 indicate the private capital and labour share in production, respectively. Equation (13) displays an additional parameter associated with the productive government capital, that is α^{kg}_1. This parameter denotes the public capital share in production. Moreover, $\varepsilon^a_{1,t}$ indicates the total factor productivity exogenous shock following a first order autoregressive process. Firms set their prices according to current and expected marginal costs, but also according to the past inflation rate. In our case, the marginal cost does not only depend on wages and the capital rental rate, but also on the price of the productive government capital.
In line with Leeper et al. (2010a) and Asimakopoulos et al. (2020), we assume that the evolution equation for productive government capital is given by:

$$K^{gp}_{1,t+1}\left(j\right)=\left(1-\delta^g_1\right)K^{gp}_{1,t}\left(j\right)+G^p_{1,t}, \tag{14}$$

where δ^g_1 is the parameter indicating the depreciation rate of the productive government capital. Moreover, $G^p_{1,t}$ indicates the productive government investment.
We also assume that intermediate production firms set prices according to the Calvo (1983) model. As an additional assumption concerning nominal rigidities, we allow for partial indexation of both wages and prices to past inflation rates.

Fiscal authority. The government finances its public spending by issuing bonds or adjusting taxes and transfers. As in Asimakopoulos et al. (2020), we separate government expenditure into productive (G_t^p) and unproductive (G_t^u). Therefore, the fiscal authority's period-by-period budget constraint has the following form:

$$P_{1,t}^{gu} G_{1,t}^{u} + P_{1,t}^{gp} G_{1,t}^{p} + B_{1,t} + T_{1,t} = \tau_{1,t}^{r} + \left(R_{1,t}^{b}\right)^{-1} B_{1,t+1},$$

where $\tau_{1,t}^r$ denotes the total government distortionary tax revenues that are given by:

$$\tau_{1,t}^{r} = \tau_{1,t}^{l} W_{1,t} L_{1,t} + \tau_{1,t}^{k} R_{1,t}^{k} K_{1,t-1}.$$

In line with Leeper et al. (2010a) and Asimakopoulos et al. (2020), we assume that the log-linearized expressions for the fiscal policy rules are:

$$\hat{\tau}_{1,t}^{l} = \phi_1^{yl} \hat{y}_{1,t}^{d} + \gamma_1^{bl} \hat{b}_{1,t-1} + \hat{\varepsilon}_{1,t}^{l} \tag{15}$$

$$\text{where} : \hat{\varepsilon}_{1,t}^{l} = \rho_1^{l} \hat{\varepsilon}_{1,t-1}^{l} + \eta_t^{l} \tag{16}$$

$$\hat{\tau}_{1,t}^{k} = \phi_1^{yk} \hat{y}_{1,t}^{d} + \gamma_1^{bk} \hat{b}_{1,t-1} + \hat{\varepsilon}_{1,t}^{k} \tag{17}$$

$$\text{where} : \hat{\varepsilon}_{1,t}^{k} = \rho_1^{k} \hat{\varepsilon}_{1,t-1}^{k} + \eta_t^{k} \tag{18}$$

$$\hat{t}_{1,t} = -\phi_1^{yt} \hat{y}_{1,t}^{d} - \gamma_1^{bt} \hat{b}_{t-1} + \hat{\varepsilon}_{1,t}^{t} \tag{19}$$

$$\text{where} : \hat{\varepsilon}_{1,t}^{t} = \rho_1^{t} \hat{\varepsilon}_{1,t-1}^{t} + \eta_t^{t} \tag{20}$$

$$\hat{g}_{1,t}^{p} = \phi_1^{ygp} \hat{y}_{1,t}^{d} - \gamma_1^{bgp} \hat{b}_{t-1} + \hat{\varepsilon}_{1,t}^{gp} \tag{21}$$

$$\text{where} : \hat{\varepsilon}_{1,t}^{gp} = \rho_1^{gp} \hat{\varepsilon}_{1,t-1}^{gp} + \eta_t^{gp} \tag{22}$$

$$\hat{g}_{1,t}^{u} = -\phi_1^{ygu} \hat{y}_{t} - \gamma_1^{bgu} \hat{b}_{1,t-1} + \hat{\varepsilon}_{1,t}^{gu} \tag{23}$$

$$\text{where} : \hat{\varepsilon}_{1,t}^{gu} = \rho_1^{gu} \hat{\varepsilon}_{1,t-1}^{gu} + \eta_t^{gu} \tag{24}$$

where the small letters with the hats denote log-deviations of the variables from their respective steady states. Moreover, we assume that all the coefficients in the fiscal rules have positive values, i.e., $\phi_1^x \geq 0$ for $x = \{yl, yk, yt, ygp, ygu\}$ and $\gamma_1^z \geq 0$ for $z = \{bl, bk, bt, bgp, bgu\}$. Fiscal rules (15)-(23) imply that fiscal variables respond to contemporaneous changes in output and with a delay of one quarter to variations in government debt. Moreover, equations (15), (17), (19), (21) and (23) include five distinct exogenous AR(1) processes , $\hat{\varepsilon}_{1,t}^{l}$, $\hat{\varepsilon}_{1,t}^{k}$, $\hat{\varepsilon}_{1,t}^{t}$, $\hat{\varepsilon}_{1,t}^{gp}$ and $\hat{\varepsilon}_{1,t}^{gu}$, with $\rho_1^v \in [0,1]$ for $v = \{l, k, t, gp, gu\}$ and each of the η_1's are distributed i.i.d. $\mathcal{N}(0,1)$. Finally, we assume that productive spending responds positively to increases in aggregate output (see, for example, Ambler et al., 2017), whereas unproductive government spending responds negatively (see, for example, Leeper et al., 2010a).

Central bank. The central bank is assumed to follow a Taylor-type interest-rate rule (Taylor, 1993) specified in terms of the past nominal interest rate, domestic inflation and output gap:[7]

$$\frac{R^b_{1,t}}{\left(R^b_1\right)^{SS}} = \left(\frac{R^b_{1,t-1}}{\left(R^b_1\right)^{SS}}\right)^{\rho^r_1} \left[\left(\frac{\pi^d_{1,t}}{\left(\pi^d_1\right)^{SS}}\right)^{r^\pi_1} \left(\frac{Y^d_{1,t}}{Y^{dp}_{1,t}}\right)^{r^y}\right]^{\left(1-\rho^r_1\right)} \left(\frac{Y^d_{1,t}/Y^d_{1,t-1}}{Y^{dp}_{1,t}/Y^{dp}_{1,t-1}}\right)^{r^{\Delta y}_1} \varepsilon^r_{1,t} \tag{25}$$

where $\left(R^b_1\right)^{SS}$ and $\left(\pi^d_1\right)^{SS}$ indicate the steady-state values for the nominal interest rate and domestic inflation, respectively. Moreover, we denote by ρ^r_1 the interest rate smoothing parameter, while r^y denotes the response of the nominal interest rate to the output gap, $r^{\Delta y}_1$ indicates the response of the nominal interest rate to changes in the output gap, and r^π_1 represents the reaction of the interest rate on domestic inflation. We denote by $\varepsilon^r_{1,t}$ the monetary policy shock that follows an AR(1) process.

Market clearing condition. Imposing the market-clearing condition for the good market of the domestic economy implies the following aggregate resource constraint:

$$Y^d_{1,t} = C^d_{1,t} + I^d_{1,t} + G^d_{1,t} + \frac{\zeta_2}{\zeta_1} M_{2,t} \tag{26}$$

$$\text{where}: M_{2,t} = M^c_{2,t} + M^i_{2,t}$$

where $M_{2,t}$ indicates the net imports of the foreign country, while ζ_1 and ζ_2 represent the relative population sizes of the home and foreign country, respectively. Simply, the market clearing condition (26) states that the production of domestic firms equals the domestic demand of the representative household for consumption and investment goods, plus domestic government expenditure and total imports of the foreign country.

Bilateral relations. For country 1, the relative import prices can be expressed as follows:

$$\frac{P^m_{1,t}}{P^d_{1,t}} = \frac{e_{1,t} P^c_{2,t}}{P^c_{1,t}} \frac{P^d_{2,t}}{P^c_{2,t}} \frac{P^c_{1,t}}{P^d_{1,t}} \tag{27}$$

[7]We define the output gap as the difference between actual ($Y^d_{1,t}$) and potential output ($Y^{dp}_{1,t}$).

where $P^m_{1,t}$ is the price of imported goods, whereas $P^d_{1,t}$ indicates the price of the final production good. Moreover, the consumption real exchange rate is given by:

$$rer_{1,t} = \frac{e_{1,t} P^c_{2,t}}{P^c_{1,t}} \tag{28}$$

We assume that the domestic holdings of internationally traded bonds (that is, the home country's net foreign assets, denominated in foreign currency) evolve according to:

$$\frac{e_{1,t} \left(R^b_{2,t}\right)^{-1} B^f_{1,t+1}}{\phi^b_{1,t}} = e_{1,t} B^f_{1,t} + \frac{\zeta_2}{\zeta_1} e_{1,t} P^m_{2,t} \left(M^c_{2,t} + M^i_{2,t}\right) - P^m_{1,t} \left(M^c_{1,t} + M^i_{1,t}\right) \tag{29}$$

where $M^c_{2,t}$ and $M^i_{2,t}$ indicate the foreign country imports of consumption and investment goods, respectively. Finally, the market clearing condition for the holdings of foreign assets states that $B^f_{1,t} + B^f_{2,t} = 0$.

3.2 Data

We estimate the model using data for Canada and the US for the sample period 1981:Q1-2019:Q1. We have chosen this pair of countries for our analysis because the trade between Canada and the US accounts for approximately 70% of total Canadian trade. This implies that the trade with the US provides a realistic characterization of the rest of the world for Canada.

Since there are twenty-two exogenous shocks in the model, twenty-two data series are used in the estimation. In particular, we use data on Canadian and US real gross domestic products, Canadian and US real private investments, Canadian and US real wage compensations, Canadian and US inflation rates, Canadian and US nominal interest rates, Canadian and US real labour tax revenues, Canadian and US real capital tax revenues, Canadian and US real productive government expenditures, Canadian and US real unproductive government expenditures, Canadian and US real government lump-sum transfers, Canadian real imports from the US and US real imports from Canada.

We use the OECD Economic Outlook no. 106 database as the primary source for most of our variables. The only exceptions are the series for the wage compensations, the nominal interest rates and the imports. The series for the wage compensation in Canada is taken from Statistics Canada. The series for the wage compensation in the US is taken from the US FRED. The series of the Canadian nominal interest rate is constructed using data from the IMF and the Bank of

Canada. The series of the US nominal interest is taken from the US FRED. The series of imports for both countries are taken from the IMF (Direction of Trade Statistics).

For each country, to obtain the real variables, we deflate the nominal variables using the country-specific GDP deflator.[8] Then, for each country, the real variables are converted into per capita terms by dividing for the country-specific working-age population.

Following Asimakopoulos et al. (2020), we assume that the productive government spending includes the expenditure with a substantial (physical or human) capital component, whereas the unproductive spending category relates to government final wage and non-wage consumption expenditures. Accordingly, government productive expenditure is composed of government fixed capital formation, capital payments and government consumption of fixed capital. Unproductive government spending corresponds to government final consumption expenditure. We also assume that the series of government transfers is given by the social security benefits paid by the government.

The detailed description of data construction and sources for the observed variables of the model are reported in online Appendix B.

3.3 Model parameters

We choose to split the parameters into three different sets: the first corresponds to parameters that are kept fixed and are set accordingly to the previous economic literature; the second set is constructed from the observed data; and the third set is estimated with the MitISEM.

Fixed and calibrated parameters according to actual data. Table 5 presents the first set of parameters which can be viewed as strict priors because they can be directly related to the steady-state values and are not identifiable from the data we use. In order to set up the values of these parameters, we follow the most recent DSGE literature. Moreover, we assume that these parameters have the same values for both domestic and foreign countries. Thus, the following parameters have subscript i, indicating $i = \{Canada, US\}$.

We fix the discount factor (β_i) in line with the value assumed by Del Negro and Schorfheide (2008). As it is common in the literature, we assume a private capital depreciation rate (δ_i) that implies an annual depreciation on capital of 0.10. We assume that the intertemporal elasticity of substitution

[8] The only exception is the series of Canadian imports from the US. Since the original series is expressed in US dollars, we use the US GDP deflator to deflate this series.

($\frac{1}{\sigma_i^c}$) corresponds to a coefficient of relative risk aversion equal to 5. This value of the risk aversion is commonly used in the macroeconomic literature (see, for example, Asimakopoulos et al., 2020 and Jermann, 1998). In line with previous macroeconomic models, we set up the elasticity of labour supply (σ_i^l) equal to 4 (for a detailed survey, see Chetty et al., 2013). As in SW, the steady-state mark-up in the labour market (ν_i^w) is equal to 1.5, and we assume that the steady-state mark-up in the goods market (ν_i^p) is equal to 1.5 as well. Moreover, the curvature parameters of the Kimball aggregators in the goods, (ϑ_i^p), and labour, (ϑ_i^w), market are both set at 10. We follow Bodenstein et al. (2011) and assume a value of 0.0001 for the parameter capturing the curvature of the bond intermediation cost (ϕ_i^b). As in Leeper et al. (2010a) and Asimakopoulos et al. (2020), we assume that the depreciation rate for the government capital expenditure (δ_i^g) corresponds to 0.005. Moreover, we assume a value of the private capital share in the production function (α_i^k) that is in line with the Leeper et al. (2010b) calibration. We set the parameter indicating the public capital share in the production function (α_i^{kg}), which is in line with the estimates of Asimakopoulos et al. (2020).

In Table 6, we report the second set of parameters that are constructed from the observed data of Canada and the US. Once these parameters are computed, we hold them as fixed in order to estimate the model. For both countries, the relative shares of productive and unproductive government expenditures on GDP are computed as average ratios for the 1981-2019 period. Similarly, in each country, the steady-state tax rates for capital and labour are obtained from average capital and labour income tax rates, respectively, taken from our sample data. In each country, the share of transfers over GDP has been computed residually from the government's budget constraint using the steady states reported above and the relative steady state of debt to output ratio, which is the average annual debt to output ratio for the period under consideration.[9]

In the period 1981-2019, the Canadian imports of goods and services from the US accounted for approximately 19% of Canadian GDP. During the same period, Canadian total imports were divided into 82% consumption goods and 18% services approximately. On the other hand, the US imports of goods and services from Canada accounted for approximately 2% of US GDP. US total imports were divided into 83% consumption goods and 17% services approximately. By combining these statistics, we are able to compute the steady-state parameters and determine trade flows for Canada and the US: the parameters measuring the weight on imports in consumption (ω_1^{mc}

[9] During the period 1981-2019, the average shares of annual debt over output in Canada and the US are approximately 55% and 66%, respectively.

and ω_2^{mc}, respectively), as well as the parameters capturing the weight on imports in investment (ω_1^{mi} and ω_2^{mi}, respectively).[10] Finally, from Table 6, we observe that the share of the Canadian population on the sum of the population of the two countries corresponds approximately to 10%.

Prior distributions. Table 7 shows the third group that includes the endogenous parameters for both Canada and the US and is estimated with the MitISEM. We choose priors that are the same for both countries and are in line with previous literature. More specifically, we set the priors for habit in consumption , investment adjustment costs and Calvo probabilities for both wages and prices in accordance with Del Negro and Schorfheide (2008). Moreover, we set the priors for indexation parameters of both wages and prices that are in line with Cacciatore and Traum (2020). Turning to the monetary policy rule, the prior for the degree of interest rate smoothing and the priors for the long-run reaction coefficients of inflation and output are in line with those used by Del Negro and Schorfheide (2008). Following Asimakopoulos et al. (2020), we set the prior of the short-run coefficient of output as Gamma distributed with mean equal to 1.20 and standard deviation of 0.05.

Focusing on the coefficients of the fiscal sector, our priors are in line with Asimakopoulos et al. (2020). More specifically, the priors for the parameters of lump-sum transfers and labour tax rate elasticities with respect to output are assumed to have Gamma distributions with the same mean of 0.10 and the same standard deviation of 0.05. In addition, we assume that the prior for the parameter of the capital tax rate elasticity with respect to output is Gamma distributed with mean of 0.40 and standard deviation 0.20. Our assumed prior distributions for the responses of labour income tax, capital tax and lump-sum transfers to government debt range approximately between 0 and 0.25, between 0 and 0.75, and between 0 and 1. We assume that the parameters that measure the responses of productive and unproductive government expenditures to output have Gamma distribution with mean of 0.15 and standard deviation of 0.05. Moreover, the priors for the parameters that indicate the responses of productive and unproductive government expenditures to debt are Gamma distributed with mean of 0.40 and standard deviation of 0.20.

In the last row of Table 7, we focus on the choice of the prior for the elasticity of substitution between domestic and foreign goods. Following Cacciatore and Traum (2020), we assume a prior gamma distribution with mean equal to 1.10 and standard deviation equal to 0.10.

Table 8 shows the priors of the parameters related to all the exogenous processes in our model. Fol-

[10] See Online Appendix A for the full derivation of composite parameters.

lowing Del Negro and Schorfheide (2008), we use beta distributions for the persistence parameters of the several shocks with prior mean values of 0.75 and prior standard deviations of 0.15. Finally, we use inverse gamma distributions for standard errors of exogenous shocks with mean equal to 0.10 and standard deviation equal to 2.00. This prior characterization is in line with Herbst and Schorfheide (2014).

3.4 Estimation results

Table 7 shows the posterior mean estimates for the endogenous parameters with the relative credible intervals for the $5th$ and $95th$ percentiles. In general, the parameters are strongly identified with tight posterior distributions.[11]

The posterior means of consumption habit for Canada (h_1) and the US (h_2) correspond to 0.32 and 0.17, respectively. These values are lower than those found by Cacciatore and Traum (2020).[12] Focusing on the investment adjustment cost, our estimated value of φ_1^i is higher than the value found by Cacciatore and Traum (2020) for Canada, whereas our posterior mean for φ_2^i is in line with the values found by Smets and Wouters (2007) and Cacciatore and Traum (2020) for the US. In terms of nominal rigidities, our estimates indicate that wages and prices are sticky in Canada and the US. In particular, the posterior mean estimates of the Calvo wage setting probabilities for Canada (ξ_1^w) and the US (ξ_2^w) are higher than our assumed priors. Our results show that the probability of optimally resetting nominal wages in Canada is approximately 0.11 and in the US is approximately 0.13. Similarly, the estimated means of the Calvo price setting probabilities for Canada (ξ_1^p) and the US (ξ_2^p) are higher than their priors. This implies that the Calvo readjustment probability for Canada is approximately 0.18, whereas for the US it is approximately 0.15.[13] Turning to the posterior estimates of wage indexations, the mean values of both Canada and the US are higher than those found by Cacciatore and Traum (2020) and Smets and Wouters (2007) for Canada and the US, respectively. Moreover, our estimated mean of ι_1^p is higher than the value found by Cacciatore and Traum (2020) for Canada, whereas the posterior mean of ι_2^w falls within the range of values found by Smets and Wouters (2007) and Del Negro and Schorfheide (2008) for the US.

[11] In online Appendix C, we show all the prior and posterior density functions for the estimated parameters.

[12] The estimated values found by Cacciatore and Traum (2020) are 0.72 and 0.84 for Canada and the US, respectively.

[13] Our estimated values of ξ_1^w and ξ_1^p are higher than the values found by Cacciatore and Traum (2020) for Canada, whereas our estimated posteriors for ξ_2^w and ξ_2^p are in line with those found by Smets and Wouters (2007) and Del Negro and Schorfheide (2008) for the US.

Focusing on our estimates of the monetary policy reaction functions, the posterior mean of the reaction coefficient to inflation is estimated to be lower in Canada than in the US. In both countries, the nominal interest rate reacts much less strongly to the output gap in the long-run than in the short-run. The posterior of the degree of interest rate smoothing is higher in the US than in Canada.

Now we focus on the estimated posteriors of the fiscal rule parameters. In both countries, we observe that the capital tax response is more procyclical than the labour tax response.[14] For both Canada and the US, capital tax responds more strongly than labour tax to changes in government debt.[15] Our estimated results are in line with many studies in the optimal fiscal policy literature for the US economy (see, for example, Barro, 1979; Chari et al., 1994; Angelopoulos et al., 2015). In addition, our estimates indicate that, in both Canada and the US, lump-sum transfers respond more strongly to changes in the debt-to-output ratio than to output deviations. This result implies that non-distortionary taxation is the preferred option to stabilize debt in both Canada and the US.

Focusing on the two different types of government expenditure, we observe important differences between Canada and the US. In particular, our estimated results show that, in Canada, the productive government spending has a stronger response to changes in output than unproductive expenditure. The opposite result is found for the US. Moreover, in Canada, productive government expenditure responds less strongly than unproductive government spending to debt variations. On the contrary, our posterior estimates for the US show that $\gamma_2^{bg^p}$ is higher than $\gamma_2^{bg^u}$.[16]

The posterior estimates of the elasticity between domestic and foreign goods for both Canada and the US are very close to unity. Accordingly, our findings are line with the estimated values by Cacciatore and Traum (2020).[17]

Turning to the exogenous shocks, Table 8 shows the estimated posteriors for the autocorrelation coefficients and standard errors of all the exogenous processes, together with their credible intervals for the $5th$ and $95th$ percentiles.

In general, all the exogenous disturbances seem to be well identified with tight posterior distributions. As concerns $AR(1)$ processes, the estimated persistence of all taxes and productive and

[14] The procyclicality of capital tax is especially strong for Canada.

[15] The estimated response of capital tax to government debt is particularly strong in Canada.

[16] Our results for the estimated parameters of the two types of spending for the US confirm the findings by Asimakopoulos et al. (2020), although with a different magnitude in the posterior means.

[17] The values of our estimated parameters are slightly lower than those found by Cacciatore and Traum (2020) for the substitutability of home and foreign between Canada and the US.

unproductive government spending are less persistent in Canada than in the US.[18] Notable differences in estimated persistence also relate to investment and productivity shocks, with the former higher in Canada whereas the latter is higher in the US.

Finally, our posterior estimates show that Canada and the US have similar estimated volatilities.[19] Focusing on Canada, we note that productivity, import preferences, investment, and unproductive and productive government expenditure shocks are the most volatile shocks. Focusing on the US, we observe that monetary policy, unproductive government spending, investment, and productive government spending are more volatile than the remaining shocks.

3.5 Impulse response analysis

Now we focus on some impulse response functions for the estimated model. In particular, we analyze the IRFs related to productive and unproductive government spending shocks for the home country, i.e., Canada. The lines displayed in the various charts are generated by the mean estimates of the posterior distributions. In each figure, we show the impulse responses following a 1% exogenous positive shock to domestic productive and unproductive government spending.[20] In Figures 3 and 4, we include two lines: i) the solid line representing the model with nominal rigidities; ii) the dashed line indicating the model with flexible prices and wages (without nominal rigidities).[21]

Productive government expenditure. Figure 3 shows that in both models, with and without nominal rigidities, a positive shock to domestic productive government spending induces an increase in domestic output for all periods. Moreover, hours worked rise in response to this shock.

Importantly, we observe that domestic private consumption behaves differently in the model with nominal rigidities with respect to the model with flexible prices and wages. In the former case, private consumption increases, whereas in the latter case it decreases. The response of private consumption depends on the reaction of the domestic wage rate. In the model with nominal rigidities, the domestic wage rate increases, whereas the opposite occurs in the economy without nominal rigidities. The intuition behind this result is the following. An increase in productive

[18] The only exception is the capital tax.

[19] The exception relates to the monetary policy shock that is estimated to be much more volatile for the US economy.

[20] Qualitatively the results of the IRFs are the same if we use the estimated standard deviation of the shocks instead of the simulated 1% standard deviation. We simply normalize the shock to the economy to be 1% to ease the comparison of the impulse responses between the two cases of domestic productive and unproductive government spending. In online Appendix D, we present the estimated impulse responses together with the confidence intervals.

[21] As in Smets and Wouters (2007), the model with flexible prices and wages is obtained by removing nominal rigidities, as well as price and wage mark-up shocks from the model with rigid prices and wages. In online Appendix A, we report the equations for the flexible-price-and-wage version of the model.

government spending leads to a rise in aggregate demand, as well as an increase in the marginal product of labour. In turn, this implies that labour demand rises. At the same time, higher future distortionary taxes imply a negative wealth effect on the households that increase their labour supply. The net effect on wages and consumption depends on whether labour demand or labour supply rises more. With nominal rigidities, firms cannot adjust prices, but have to satisfy higher demand, so they increase their labour demand by more than in the model with no nominal rigidities.

In the model with nominal rigidities, the increase in domestic productive government spending induces an increase in the domestic firms' marginal cost and inflation. In particular, we note that the ratio of the price over the marginal cost decreases following the productive spending shock because marginal cost increases more than inflation. This effect, in turn, puts an additional upward pressure on prices. The central bank responds to the increase in inflation by raising the nominal interest rate.

Focusing on the domestic fiscal sector, we note that both capital and labour taxes increase in response to the shock, whereas government transfers decline. However, these effects only partially fund the exogenous increase in public spending and, as a consequence, public debt increases.

Turning to the external sector, the real exchange rate appreciates in response to the productive government shock, whereas net exports decline on impact, but then they increase over time. As argued by Born et al. (2013), previous literature has not found an agreement about the sign of the response of the trade balance to government spending shocks. Kim and Roubini (2008), Müller (2008) and Ilzetzki et al. (2013) found an improvement in the trade balance in response to expansionary fiscal shocks. On the other hand, Monacelli and Perotti (2010) and Corsetti et al. (2012) find that expansionary fiscal measures worsen the trade balance.

Our findings are in line with many theoretical models that are able to generate the crowding-in effect in response to an increase in government spending (see, for example, Baxter and King, 1993; Ambler and Paquet, 1996; Linnemann and Schabert, 2006; Ravn et al., 2006; Galí et al., 2007; Bilbiie et al., 2008; Linnemann, 2006; Bilbiie, 2009; Bilbiie, 2011; Coenen et al., 2012). More specifically, we extend the results of Asimakopoulos et al. (2020) in an open economy context, who found that, in a closed economy model, the two key elements that lead to a positive reaction in private consumption are an increase in productive spending, as well as nominal rigidities.

Moreover, our results confirm the findings by most standard models, i.e., that higher government spending appreciates the real exchange rate. According to the Mundell-Fleming model, under a flexible exchange rate regime, an increase in domestic government spending induces the domestic nominal interest rate to rise. In turn, a higher domestic nominal interest attracts capital inflows, and the domestic currency appreciates. In line with these predictions, Backus et al. (1994) used a two-country real business cycle model and found an appreciation of the real exchange rate in response to a positive government spending shock. Monacelli and Perotti (2010) developed a small open-economy new Keynesian model and showed that higher government spending appreciates the real exchange rate.

Unproductive government expenditure. As we can see from Figure 4, following the domestic unproductive government spending shock, the reactions of domestic output and hours worked are again positive.

We also note that domestic unproductive public spending exhibits higher persistence compared to domestic productive public spending. This causes a different reaction in the fiscal rules with domestic labour and capital taxes that remain high for a longer period, whereas lump-sum transfers keep decreasing over time. As a consequence, the response of domestic debt is much higher compared to the case of productive spending.

Importantly, we observe that the response of domestic real wage is negative, inducing a crowding-out effect on domestic consumption. This is the classical crowding-out effect on private consumption because the persistent high taxes cause a significant negative wealth effect on consumers.

The responses of domestic firms' marginal cost, inflation and nominal interest rate have similar patterns to those observed in response to a productive government spending shock, yet stronger magnitude. Focusing on the external sector, the real exchange rate appreciates, and the trade balance improves for all the periods considered.

To conclude, our results show that unproductive government spending has a negative effect on private consumption. This confirms the prediction of standard RBC models, in which an increase in government spending lowers the present value of after-tax income, thus generating a negative wealth effect that induces a cut in consumption (see, for example Aiyagari et al., 1992; Baxter and King, 1993; Christiano and Eichenbaum, 1992). In this regard, we extend the findings of Asimakopoulos et al. (2020) to an open economy setting, who found that, in a closed economy model, an increase in unproductive government spending induces a crowding-out effect on private consumption even in the presence of nominal rigidities. Finally, our results contribute to previous literature by showing that, regardless of the component of the public spending that increases, there is a real appreciation of the exchange rate.

3.6 Government spending multipliers

Now we turn to the analysis of multipliers for the two types of government spending. We construct the present-value multipliers following Leeper et al. (2010b). In particular, we have that:

$$\frac{\sum_{i=0}^{k}\left(\prod_{j=0}^{i} r_{1,t+j}^{-1}\right)\Delta X_{1,t+i}^{d}}{\sum_{i=0}^{k}\left(\prod_{j=0}^{i} r_{1,t+j}^{-1}\right)\Delta G_{1,t+i}^{c}} \tag{19}$$

where $X_{1,t+i}^{d}$ can denote either domestic output ($Y_{1,t+i}^{d}$) or domestic consumption ($C_{1,t+i}^{d}$). Moreover, $G_{1,t+i}^{c}$ can represent either productive ($G_{1,t+i}^{p}$) or unproductive government spending ($G_{1,t+i}^{u}$). In equation (19), $\Delta X_{1,t+i}^{d}$ and $\Delta G_{1,t+i}^{c}$ indicate the relative level changes of the variables with respect to their steady-state values. Finally, the discount factor (r_1) represents the real interest rate for the domestic economy.

Table 9 shows the cumulative present-value multipliers for output and consumption based on the mean estimates of our estimated model. The parameter k determines the period in quarters. We present the results on the impact of the exogenous shock, together with the results for 1 year ahead, 3 years ahead and for the infinite horizon case ($k = 1,000$). In addition, we compute the minimum and the maximum value of the respective multipliers.

Our results indicate that output multipliers for the domestic economy are larger in the presence of nominal rigidities than when prices and wages are fully flexible. We find present-value multipliers for output that are within the range of estimated values by Owyang et al. (2013) for the Canadian economy. We contribute with respect to previous literature by showing that output present-value multipliers are higher in response to productive government spending shocks than unproductive government expenditure shocks. In the case of the economy without rigidities to prices and wages, we find present-value multipliers for output that are in line with previous empirical studies on OECD countries (see, for example, Corsetti et al., 2012 and Alesina et al., 2018).

We also investigate consumption present-value multipliers for the domestic economy. In the presence of nominal rigidities, these multipliers assume opposite values in response to productive and

unproductive government spending shocks. In the case of a positive shock to productive government expenditure, present-value multipliers for consumption are positive both in short-run and in the long-run. On the other hand, in response to a positive shock to unproductive government spending, present-value multipliers for consumption assume negative values for all the period considered.
Turning to the case of the economy with flexible prices and wages, consumption present-value multipliers have negative values irrespective to type of government spending shock. Since firms can immediately adjust their prices in response to the government spending shock, the shift in labour demand is weaker than the one in labour supply. Accordingly, the wage rate falls. In this case, a positive shock to productive spending is not enough to induce an increase in domestic consumption. Finally, we observe a stronger fall in consumption in response to an unproductive government spending shock.

The role of trade openness following productive and unproductive government spending shocks. As a robustness exercise, we assess how an increase in both productive and unproductive government spending affect output and consumption present-value multipliers, assuming different values for the prior of the elasticity of substitution between domestic and foreign goods ($\frac{1+\rho^x}{\rho^x}$, where $x = c, i$).
Table 10 reports the comparison between the cumulative present-value multipliers of the benchmark model with a model that has a high elasticity of substitution between domestic and foreign goods (implying a larger trade openness) and a model that presents a low elasticity of substitution between domestic and foreign goods (implying a smaller trade openness).[22]
Overall, our results show that the degree of trade openness plays an important role in the transmission of productive and unproductive government spending shocks on consumption and output. Focusing on the economy with nominal rigidities, we observe that output and consumption present-value multipliers after one and three years are larger when the degree of trade openness is higher. On the contrary, with a low elasticity of substitution between domestic and foreign goods we have smaller values for output and consumption present-value multipliers after one and three years. These results are consistent with both types of government spending shocks.
Turning to the economy with flexible prices and wages, we observe that present-value multipliers show striking differences with respect to the previous case. In particular, we note that in response

[22] In the benchmark model, $\frac{1+\rho^x}{\rho^x}$ is equal to its estimated value. In the model with larger trade openness we assume $\frac{1+\rho^x}{\rho^x} = 2.43$, whereas in the model with smaller trade openness we assume $\frac{1+\rho^x}{\rho^x} = 1.00$.

to a positive government spending shock, a larger degree of openness implies lower output and consumption present-value multipliers. Again this outcome is consistent with each type of government spending shock.

Accordingly, in line with Cacciatore and Traum (2020), our results indicate that the degree of trade openness matters in terms of propagation of government spending shocks. In particular, when the economy presents nominal rigidities, we observe higher output and consumption multipliers when the elasticity of substitution between domestic and foreign goods is larger. The opposite is true in the case of the economy with flexible prices and wages.

4 Conclusions

In this paper, we introduced a new adaptive methodology for DSGE estimation. Our estimation procedure is based on the Mixture of (Student's) t by Importance Sampling weighted Expectation Maximization. The algorithm applies IS to compute the unknown posterior density by a mixture of Student's t densities and adapts parameters to the most recent IS draws via an EM step. The algorithm can be easily parallelized and does not require tuning parameters for the user. Moreover, it is flexible and can handle different unknown and complex forms in a reasonable computing time. We performed simulation exercises on two workhorse DSGE models, namely the small new Keynesian (NK) model (Woodford, 2003) and the Smets and Wouters (2007) (SW) model. Our results provide evidence on how the MitISEM achieves identification of posterior distributions in several cases including bimodality.

We also applied the MitISEM in order to estimate an open economy DSGE model encompassing international trade between Canada and the US. Our findings show that, in the presence of nominal rigidities, an increase in productive spending generates a crowding-in on domestic private consumption, whereas unproductive spending induces a fall in domestic private consumption. Moreover, we find that irrespective of the type of government expenditure, an increase in public spending for the domestic economy induces an exchange rate appreciation and an improvement in the trade balance. We also analysed present-value multipliers for domestic consumption and output. Such multipliers are larger in the presence of nominal rigidities than when prices and wages are fully flexible. Moreover, our findings indicate that consumption multipliers are positive only in the presence of nominal rigidities and in response to productive government spending shocks. We also show that the degree of trade openness matters in terms of transmission channels of productive and productive government spending shocks on the economy. With nominal rigidities, we have higher output and consumption multipliers when the degree of openness is larger. The opposite result is found in the case of the economy with flexible prices and wages.

References

Adjemian, S., Bastani, H., Juillard, M., Karamé, F., Maih, J., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M., and Villemot, S. (2011). Dynare: Reference Manual Version 4. Dynare Working Papers 1, CEPREMAP.

Adolfson, M., Laséen, S., Lindé, J., and Villani, M. (2007). Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through. *Journal of international economics*, 72:481–511.

Adolfson, M., Laséen, S., Lindé, J., and Villani, M. (2008). Evaluating an Estimated New Keynesian Small Open Economy Model. *Journal of Economic Dynamics and Control*, 32(8):2690–2721.

Aiyagari, S. R., Christiano, L. J., and Eichenbaum, M. (1992). The Output, Employment, and Interest Rate Effects of Government Consumption. *Journal of Monetary Economics*, 30:73–86.

Alesina, A., Favero, C. A., and Giavazzi, F. (2018). What Do We Know about the Effects of Austerity? In *AEA Papers and Proceedings*, volume 108, pages 524–30.

Ambler, S., Bouakez, H., and Cardia, E. (2017). Does the Crowding-in Effect of Public Spending on Private Consumption Undermine Neoclassical Models? *Research in Economics*, 71:399–410.

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

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE models. *Econometric Reviews*, 26:113–172.

Andrews, I. and Mikusheva, A. (2015). Maximum Likelihood Inference in Weakly Identified Dynamic Stochastic General Equilibrium Models. *Quantitative Economics*, 6:123–152.

Angelopoulos, K., Asimakopoulos, S., and Malley, J. (2015). Tax Smoothing in a Business Cycle Model with Capital-Skill Complementarity. *Journal of Economic Dynamics and Control*, 51:420–444.

Asimakopoulos, S., Lorusso, M., and Pieroni, L. (2020). Can Public Spending Boost Private Consumption? *Canadian Journal of Economics*, Forthcoming.

Auerbach, A. J. and Gorodnichenko, Y. (2016). Effects of Fiscal Shocks in a Globalized World. *IMF Economic Review*, 64:177–215.

Backus, D. K., Kehoe, P. J., and Kydland, F. E. (1994). Dynamics of the Trade Balance and the Terms of Trade: The J-Curve? *The American Economic Review*, pages 84–103.

Baştürk, N., Grassi, S., Hoogerheide, L., Opschoor, A., and van Dijk, H. (2017). The R Package MitISEM: Efficient and Robust Simulation Procedures for Bayesian Inference. *Journal of Statistical Software, Articles*, 79:1–40.

Baştürk, N., Grassi, S., Hoogerheide, L., and Van Dijk, H. K. (2016). Parallelization Experience with Four Canonical Econometric Models Using ParMitISEM. *Econometrics*, 4:1–11.

Barro, R. J. (1979). On the Determination of the Public Debt. *Journal of Political Economy*, 87:940–71.

Baxter, M. and King, R. G. (1993). Fiscal Policy in General Equilibrium. *The American Economic Review*, pages 315–334.

Beetsma, R., Giuliodori, M., and Klaassen, F. (2008). The Effects of Public Spending Shocks on Trade Balances and Budget Deficits in the European Union. *Journal of the European Economic Association*, 6:414–423.

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

Bilbiie, F. (2011). Nonseparable Preferences, Frisch Labor Supply, and the Consumption Multiplier of Government Spending: One Solution to a Fiscal Policy Puzzle. *Journal of Money, Credit and Banking*, 43:221–251.

Bilbiie, F., Meier, A., and Muller, G. (2008). What Accounts for the Changes in US Fiscal Policy Transmission? *Journal of Money, Credit and Banking*, 40:1439–1470.

Blanchard, O. (2017). On the Need for (at Least) Five Classes of Macro Models. *Peterson Institute for International Economics, Washington, DC.*

Blanchard, O. and Perotti, R. (2002). An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. *The Quarterly Journal of Economics*, 117:1329–1368.

Bodenstein, M., Erceg, C. J., and Guerrieri, L. (2011). Oil Shocks and External Adjustment. *Journal of International Economics*, 83:168–184.

Bognanni, M. and Herbst, E. (2018). A Sequential Monte Carlo Approach to Inference in Multiple-Equation Markov-Switching Models. *Journal of Applied Econometrics*, 33:126–140.

Born, B., Juessen, F., and Müller, G. J. (2013). Exchange Rate Regimes and Fiscal Multipliers. *Journal of Economic Dynamics and Control*, 37:446–465.

Bouakez, H., Chihi, F., and Normandin, M. (2014). Measuring the effects of fiscal policy. *Journal of Economic Dynamics and Control*, 47(C):123–151.

Cacciatore, M. and Traum, N. (2020). Trade Flows and Fiscal Multipliers. *Review of Economics and Statistics*, pages 1–44.

Cai, M., Del Negro, M., Herbst, E., Matlin, E., Sarfati, R., and Schorfheide, F. (2020). Online Estimation of DSGE Models. *The Econometrics Journal*, pages 1–30.

Calvo, G. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12:383–398.

Campbell, J. Y. (1994). Inspecting the mechanism: An analytical approach to the stochastic growth model. *Journal of Monetary Economics*, 33:463–506.

Cappé, O., Douc, R., Guillin, A., Marin, J. M., and Robert, C. P. (2008). Adaptive Importance Sampling in General Mixture Classes. *Statistics and Computing*, 18:447–459.

Chari, V., Christiano, L., and Kehoe, P. (1994). Optimal Fiscal Policy in a Business Cycle Model. *Journal of Political Economy*, 102:617–652.

Chetty, R., Guren, A., Manoli, D., and Weber, A. (2013). Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities. *NBER Macroeconomics Annual*, 27:1–56.

Chib, S. and Ramamurthy, S. (2010). Tailored Randomized Block MCMC Methods with Application to DSGE models. *Journal of Econometrics*, 155:19–38.

Christiano, L. J. and Eichenbaum, M. (1992). Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations. *The American Economic Review*, 82:430–450.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of political Economy*, 113:1–45.

Clarida, R., Gali, J., and Gertler, M. (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of economic literature*, 37:1661–1707.

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

Corsetti, G., Meier, A., and Müller, G. J. (2012). What Determines Government Spending Multipliers? *Economic Policy*, 27:521–565.

Corsetti, G. and Müller, G. J. (2006). Twin Deficits: Squaring Theory, Evidence and Common Sense. *Economic Policy*, 21:598–638.

Corsetti, G. and Pesenti, P. (2001). Welfare and Macroeconomic Interdependence. *The Quarterly Journal of Economics*, 116:421–445.

Creal, D. (2007). Sequential Monte Carlo Samplers for Bayesian DSGE Models. Working papers, Vrije Universitiet.

Del Negro, M. and Schorfheide, F. (2008). Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics*, 55:1191–1208.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society B (Methodological)*, 39:1–38.

Durham, G. and Geweke, J. (2011). Massively parallel sequential monte carlo for bayesian inference. Manuscript.

Enders, Z., Müller, G. J., and Scholl, A. (2011). How Do Fiscal and Technology Shocks Affect Real Exchange Rates?: New Evidence for the United States. *Journal of International Economics*, 83:53–69.

Erceg, C. J., Guerrieri, L., and Gust, C. (2008). Trade Adjustment and the Composition of Trade. *Journal of Economic Dynamics and Control*, 32:2622–2650.

Fatás, A. and Mihov, I. (2003). The Case for Restricting Fiscal Policy Discretion. *The Quarterly Journal of Economics*, 118:1419–1447.

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

Galí, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications.* Princeton University Press.

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

Galí, J. and Monacelli, T. (2005). Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *The Review of Economic Studies*, 72:707–734.

Ghironi, F. (2017). On the Blanchard Classification of Macroeconomic Models.

Ghironi, F. and Melits, M. J. (2005). International Trade and Macroeconomic Dynamics with Heterogeneous Firms. *The Quarterly Journal of Economics*, 120:865–915.

Herbst, E. (2012). Gradient and Hessian-based MCMC for DSGE models. Unpublished manuscript, Federal Reserve Board.

Herbst, E. and Schorfheide, F. (2014). Sequential Monte Carlo Sampling for DSGE Models. *Journal of Applied Econometrics*, 29:1073–1098.

Herbst, E. and Schorfheide, F. (2016). Bayesian Estimation of DSGE models. *Princeton Press.*

Hoogerheide, L., Opschoor, A., and Van Dijk, H. K. (2012). A Class of Adaptive Importance Sampling Weighted EM Algorithms for Efficient and Robust Posterior and Predictive Simulation. *Journal of Econometrics*, 171:101–120.

Ilzetzki, E., Mendoza, E. G., and Végh, C. A. (2013). How Big (Small?) Are Fiscal Multipliers? *Journal of monetary economics*, 60:239–254.

Jermann, U. (1998). Asset Pricing in Production Economies. *Journal of Monetary Economics*, 41:257–275.

Justiniano, A. and Preston, B. (2010). Monetary Policy and Uncertainty in an Empirical Small Open-Economy Model. *Journal of Applied Econometrics*, 25:93–128.

Kim, S. and Roubini, N. (2008). Twin Deficit or Twin Divergence? Fiscal Policy, Current Account, and Real Exchange Rate in the US. *Journal of international Economics*, 74:362–383.

Kollmann, R. (2001). The Exchange Rate in a Dynamic-Optimizing Business Cycle Model with Nominal Rigidities: A Quantitative Investigation. *Journal of International Economics*, 55:243–262.

Kydland, F. E. and Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica*, 50:1345–1370.

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

Leeper, E., Walker, T., and Yang, S.-C. (2010b). Government Investment and Fiscal Stimulus. *Journal of Monetary Economics*, 57:1000–1012.

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

Linnemann, L. and Schabert, A. (2006). Productive Government Expenditure in Monetary Business Cycle Models. *Scottish Journal of Political Economy*, 53:28–46.

Long, J. B. and Plosser, C. I. (1983). Real Business Cycles. *Journal of political Economy*, 91:39–69.

Lubik, T. and Schorfheide, F. (2005). A Bayesian Look at New Open Economy Macroeconomics. *NBER macroeconomics annual*, 20:313–366.

Monacelli, T. and Perotti, R. (2010). Fiscal Policy, the Real Exchange Rate and Traded Goods. *The Economic Journal*, 120:437–461.

Mountford, A. and Uhlig, H. (2009). What Are the Effects of Fiscal Policy Shocks? *Journal of applied econometrics*, 24:960–992.

Müller, G. J. (2008). Understanding the Dynamic Effects of Government Spending on Foreign Trade. *Journal of international money and finance*, 27:345–371.

Otrok, C. (2001). On Measuring the Welfare Cost of Business Cycles. *Journal of Monetary Economics*, 47:61–92.

Owyang, M., Ramey, V., and Zubairy, S. (2013). Are Government Spending Multipliers Greater During Periods of Slack? Evidence from Twentieth-Century Historical Data. *American Economic Review*, 103(3):129–34.

Perotti, R. (2014). Defense Government Spending Is Contractionary, Civilian Government Spending Is Expansionary. NBER Working Papers, National Bureau of Economic Research, Inc.

Rabanal, P. and Tuesta, V. (2010). Euro-Dollar Real Exchange Rate Dynamics in an Estimated Two-Country Model: An Assessment. *Journal of Economic Dynamics and Control*, 34:780–797.

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

Robert, C. and Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer Verlag, New York.

Schmitt-Grohé, S. and Uribe, M. (2004). Optimal Operational Monetary Policy in the Christiano-Eichenbaum-Evans Model of the US Business Cycle. Technical report, National Bureau of Economic Research.

Schorfheide, F. (2000). Loss Function-Based Evaluation of DSGE models. *Journal of Applied Econometrics*, 15:645–670.

Smets, F. and Wouters, R. (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European economic association*, 1:1123–1175.

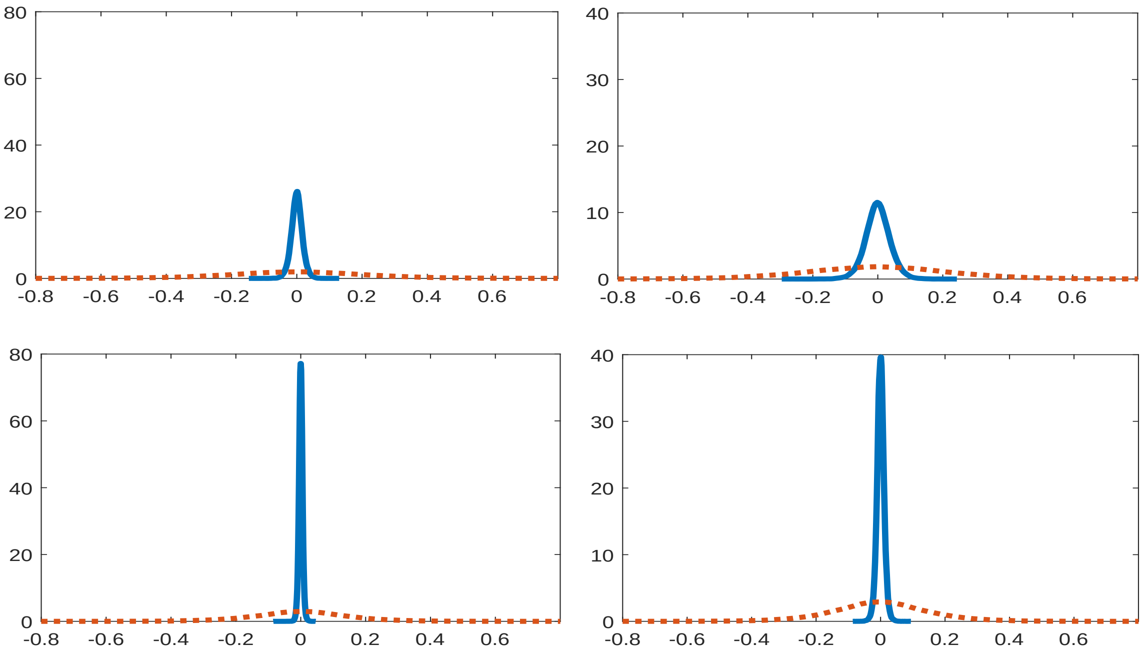
Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American economic review*, 97:586–606.

Taylor, J. B. (1993). Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214.

Woodford, M. (2003). Interest and Prices. *Princeton University Press.*

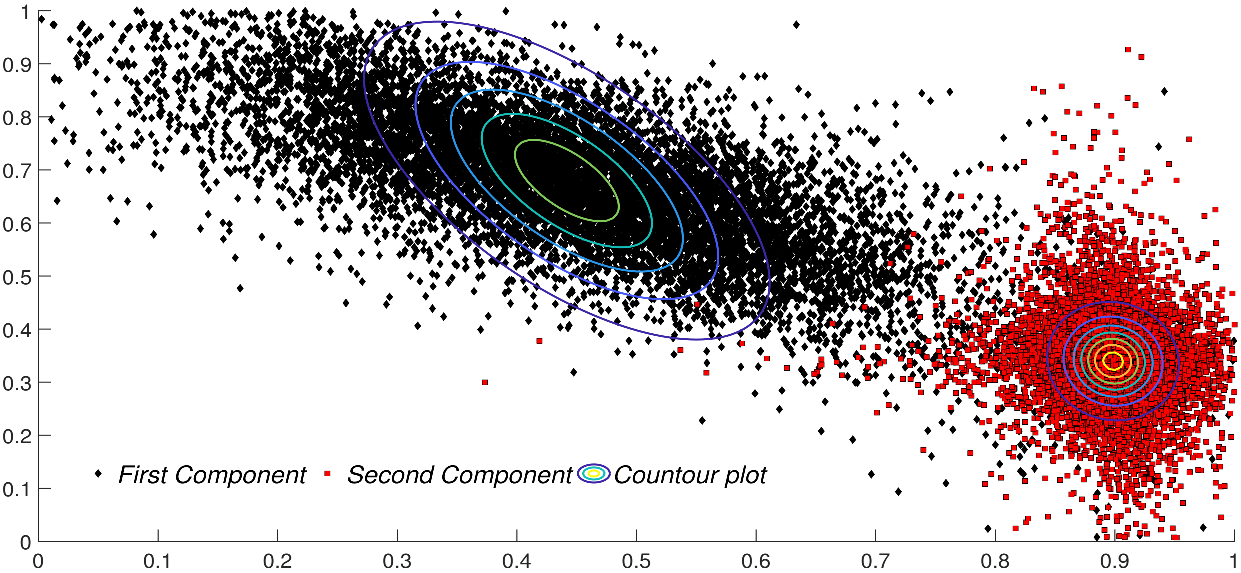
Yun, T. (1996). Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles. *Journal of Monetary Economics*, 37:345–370.

Figure 1: Posterior density of β and π



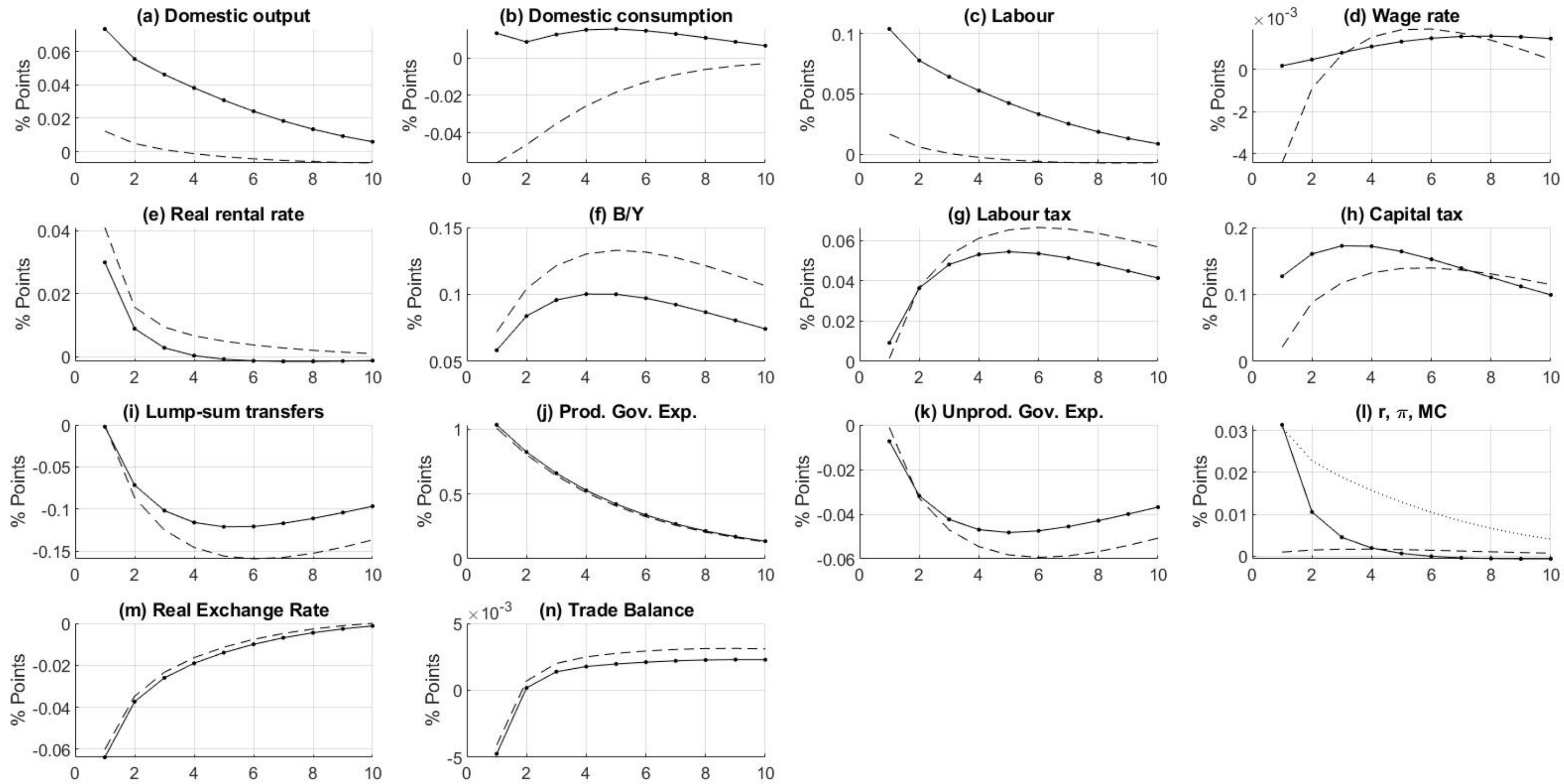
Notes: The Figure shows the posterior densities of the parameters β (left panels) and π (right panels) centred at zero to facilitate graphical comparison. The blue (solid) lines show posterior densities when the parameter π is identified; red (dotted) lines show posterior densities when the parameter is not identified. Top panels: $T = 250$; bottom panels: $T = 1000$.

Figure 2: Contour of Student's t components



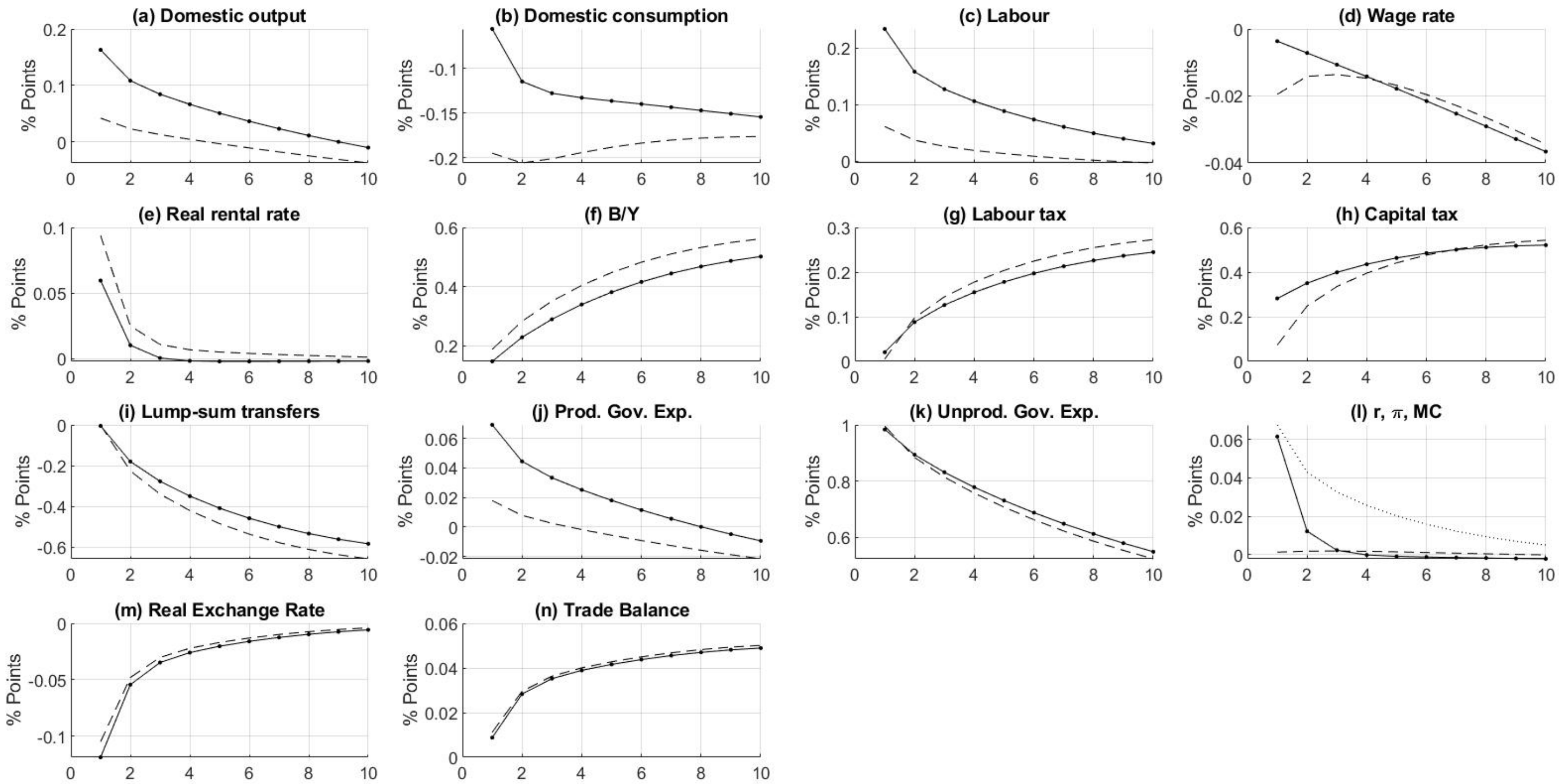
Notes: The Figure shows the contour plot of two t-student components for the parameters. The first density (black diamond) represent the $\boldsymbol{\Theta} = [0.45, 0.45]$, the parameter simulated in the experiment. The second density (red squared) represents the $\boldsymbol{\Theta} = [0.89, 0.22]$. In this experiment the number of draws is set to 100,000.

Figure 3: *Impulse responses to a 1% increase in productive public spending for the Canadian economy*



Notes: Solid lines indicate an economy with nominal rigidities. Dashed lines are for the economy without nominal rigidities. All variables are in percentage deviations from their steady state. X-axis is in quarters.

Figure 4: *Impulse responses to a 1% increase in unproductive public spending for the Canadian economy*



Notes: Solid lines indicate an economy with nominal rigidities. Dashed lines are for the economy without nominal rigidities. All variables are in percentage deviations from their steady state. X-axis is in quarters.

Table 1: *Monte Carlo exercise for the NK model*

Full name	Symbol	Prior	Mean, St.Dev.	B-MitISEM	B-MCMC
Risk aversion parameter	τ	$\mathcal{G}$	(2.00, 0.50)	0,01	0.16
Phillips curve slope coefficient	κ	U	(0.00, 1.00)	0.02	0.11
T.R. coefficient. on inflation	ψ_1	$\mathcal{G}$	(1.50, 0.25)	0.01	0.23
T.R. coefficient. on output	ψ_2	$\mathcal{G}$	(0.50, 0.25)	-0.01	0.04
Steady-state real interest rate	$r^{(A)}$	$\mathcal{G}$	(0.50, 0.50)	-0.03	0.06
Steady-state real inflation rate	$\pi^{(A)}$	$\mathcal{G}$	(7.00, 2.00)	0.03	0.08
Growth rate of the economy	$\gamma^{(Q)}$	$\mathcal{N}$	(0.40, 0.20)	0.01	0.03
Interest Rate Smoothing	ρ_r	U	(0.00, 1.00)	-0.01	0.05
Government expenditure pers.	ρ_g	U	(0.00, 1.00)	0.01	0.05
Productivity persistence	ρ_z	U	(0.00, 1.00)	0.02	0.05
Monetary policy standard error	σ_r	$\mathcal{IG}$	(0.40, 4.00)	0.00	-0.03
Government expenditure s.e.	σ_g	$\mathcal{IG}$	(1.00, 4.00)	-0.00	-0.02
Productivity standard error	σ_z	$\mathcal{IG}$	(0.50, 4.00)	-0.04	-0.07

Notes: The table shows the estimated bias for the MitISEM (B-MitISEM) and MCMC (B-MCMC). We also report the parameter's name (Full name), the acronym symbol (Symbol), and the prior distribution (Prior) with the mean and standard deviation (Mean, St.Dev). The number of Monte Carlo iterations is set to $I = 100$ and the number of draws is set to $D = 200000$. $\mathcal{G}$ is Gamma distribution. U is the Uniform distribution. $\mathcal{IG}$ is the Inverse-Gamma distribution. $\mathcal{N}$ is the Normal distribution.

Table 2: *Monte Carlo exercise for the SW model*

Full name	Symbol	Prior	Mean, St.Dev.	B-MitISEM	B-MCMC
Adjustment cost	φ	$\mathcal{N}$	(4.00, 1.50)	-0,15	-0,15
IES	σ_c	$\mathcal{N}$	(1.50, 0.37)	0,02	-0.02
Consumption habit	h	$\mathcal{B}$	(0.70, 0.10)	0.02	-0.01
Calvo wages probab.	ξ_w	$\mathcal{B}$	(0.50, 0.10)	0.02	0.00
Elasticity of labour supply	σ_l	$\mathcal{N}$	(2.00, 0.75)	0.17	0.36
Calvo prices probab.	ξ_p	$\mathcal{B}$	(0.50, 0.10)	-0.05	-0.03
Degree of wage indexation	ι_w	$\mathcal{B}$	(0.50, 0.15)	-0.03	0.00
Degree of price indexation	ι_p	$\mathcal{B}$	(0.50, 0.15)	-0.02	0.01
Capacity utilization elasticity	ψ	$\mathcal{B}$	(0.50, 0.15)	0.01	-0.01
Share of fix. costs in prod. func.	Φ	$\mathcal{N}$	(1.25, 0.12)	0.05	0.00
T.R. coefficient. on inflation	r_π	$\mathcal{N}$	(1.50, 0.25)	0.03	-0.01
Interest Rate Smoothing	ρ	$\mathcal{B}$	(0.75, 0.10)	0.03	-0.01
T.R.long-run coefficient on Y	r_y	$\mathcal{N}$	(1.50, 0.37)	0.00	0.00
T.R. short-run coefficient on Y	r_{Δ_y}	$\mathcal{N}$	(0.12, 0.05)	0.00	0.01
Steady-state inflation rate	$\bar{\pi}$	$\mathcal{G}$	(0.62, 0.10)	0.06	0.00
Steady-state discount rate	$100\left(\beta^{-1}-1\right)$	$\mathcal{G}$	(0.25, 0.10)	0.20	0.28
Steady-state hours worked	$\bar{l}$	$\mathcal{N}$	(0.00, 2.00)	-0.09	0.10
Quarterly trend growth rate	$\bar{\gamma}$	$\mathcal{N}$	(0.40, 0.10)	0.06	0.04
Private capital share	α	$\mathcal{N}$	(0.30, 0.05)	0.15	0.13
Productivity persistence	ρ_a	$\mathcal{B}$	(0.50, 0.20)	0.01	0.00
Risk-premium persistence	ρ_b	$\mathcal{B}$	(0.50, 0.20)	0.67	0.66
Government expenditure pers.	ρ_g	$\mathcal{B}$	(0.50, 0.20)	-0.05	-0.06
Investment persistence	ρ_i	$\mathcal{B}$	(0.50, 0.20)	0.01	-0.02
Monetary policy persistence	ρ_r	$\mathcal{B}$	(0.50, 0.20)	0.01	0.00
Price mark-up persistence	ρ_p	$\mathcal{B}$	(0.50, 0.20)	-0.02	-0.02
Wage mark-up persistence	ρ_w	$\mathcal{B}$	(0.50, 0.20)	0.03	0.04
MA parameter price mark-up	μ_p	$\mathcal{B}$	(0.50, 0.20)	0.04	0.01
MA parameter wage mark-up	μ_w	$\mathcal{B}$	(0.50, 0.20)	0.01	0.03
Productivity standard error	σ_a	$\mathcal{IG}$	(0.10, 2.00)	0.19	0.18
Risk-premium standard error	σ_b	$\mathcal{IG}$	(0.10, 2.00)	0.29	0.29
Government expenditure s.e.	σ_g	$\mathcal{IG}$	(0.10, 2.00)	0.04	0.05
Investment standard error	σ_i	$\mathcal{IG}$	(0.10, 2.00)	-0.18	-0.19
Monetary policy standard error	σ_r	$\mathcal{IG}$	(0.10, 2.00)	0.04	0.02
Price mark-up standard error	σ_p	$\mathcal{IG}$	(0.10, 2.00)	-0.04	-0.04
Wage mark-up standard error	$sigma_w$	$\mathcal{IG}$	(0.10, 2.00)	-0.05	0.03

Notes: The table shows the estimated bias for the MitISEM (B-MitISEM) and MCMC (B-MCMC). We also report the parameter's name (Full name), the acronym symbol (Symbol), and the prior distribution (Prior) with the mean and standard deviation (Mean, St.Dev). The number of Monte Carlo iterations is set to $I = 100$ and the number of draws is set to $D = 200000$. $\mathcal{B}$ is the Beta distribution. $\mathcal{G}$ is Gamma distribution. U is the Uniform distribution. $\mathcal{IG}$ is the Inverse-Gamma distribution. $\mathcal{N}$ is the Normal distribution.

Table 3: *Computational time needed to estimate the NK and SW models*

	NK Model			SW Model		
Number of Mixtures	*Six Threads*	*Twelve Threads*	*Twenty-four Threads*	*Six Threads*	*Twelve Threads*	*Twenty-four Threads*
2	0.84	0.50	0.44	0.61	0.31	0.23
3	1.19	0.70	0.64	0.80	0.50	0.40

Notes: The Table shows a comparison of the computational time between the MitISEM and MCMC algorithms for the NK and the SW models. The values represent the ratios in the computing time of the MitISEM with respect to the MCMC. The MitISEM is estimated by selecting 2 and 3 mixture components as well as six, twelve or twenty-four threads using a AMD Ryzen 3900X processor with 12 cores and 24 threads.

Table 4: Robustness to the initial conditions, Monte Carlo exercise for the SW model

Full name	Symbol	Prior	Mean, St.Dev.	M.D.	S.D.D.
Adjustment cost	φ	$\mathcal{N}$	(4.00, 1.50)	0.28	-0.15
IES	σ_c	$\mathcal{N}$	(1.50, 0.37)	-0.02	-0.10
Consumption habit	h	$\mathcal{B}$	(0.70, 0.10)	0.03	0.03
Calvo wages probability	ξ_w	$\mathcal{B}$	(0.50, 0.10)	0.13	0.03
Elasticity of labour supply	σ_l	$\mathcal{N}$	(2.00, 0.75)	0.15	-0.13
Calvo prices probability	ξ_p	$\mathcal{B}$	(0.50, 0.10)	0.13	0.03
Degree of wage indexation	ι_w	$\mathcal{B}$	(0.50, 0.15)	0.14	-0.03
Degree of price indexation	ι_p	$\mathcal{B}$	(0.50, 0.15)	0.10	0.02
Capacity utilization elasticity	ψ	$\mathcal{B}$	(0.50, 0.15)	-0.15	-0.04
Share of fixed costs in prod. func.	Φ	$\mathcal{N}$	(1.25, 0.12)	0.04	-0.03
T.R. coefficient. on inflation	r_π	$\mathcal{N}$	(1.50, 0.25)	-0.06	-0.04
Interest Rate Smoothing	ρ	$\mathcal{B}$	(0.75, 0.10)	0.01	0.03
T.R.long-run coefficient on Y	r_y	$\mathcal{N}$	(1.50, 0.37)	-0.05	0.00
T.R. short-run coefficient on Y	r_{Δ_y}	$\mathcal{N}$	(0.12, 0.05)	-0.08	0.00
Steady-state inflation rate	$\bar{\pi}$	$\mathcal{G}$	(0.62, 0.10)	-0.03	0.12
Steady-state discount rate	$100\left(\beta^{-1}-1\right)$	$\mathcal{G}$	(0.25, 0.10)	-0.10	-0.02
Steady-state hours worked	$\bar{l}$	$\mathcal{N}$	(0.00, 2.00)	0.13	0.13
Quarterly trend growth rate	$\bar{\gamma}$	$\mathcal{N}$	(0.40, 0.10)	0.01	0.13
Private capital share	α	$\mathcal{N}$	(0.30, 0.05)	0.12	0.02
Productivity persistence	ρ_a	$\mathcal{B}$	(0.50, 0.20)	0.03	0.03
Risk-premium persistence	ρ_b	$\mathcal{B}$	(0.50, 0.20)	-0.16	-0.01
Government expenditure persistence	ρ_g	$\mathcal{B}$	(0.50, 0.20)	0.03	-0.01
Investment persistence	ρ_i	$\mathcal{B}$	(0.50, 0.20)	0.09	-0.06
Monetary policy persistence	ρ_r	$\mathcal{B}$	(0.50, 0.20)	0.03	-0.07
Price mark-up persistence	ρ_p	$\mathcal{B}$	(0.50, 0.20)	0.02	0.01
Wage mark-up persistence	ρ_w	$\mathcal{B}$	(0.50, 0.20)	-0.05	-0.03
MA parameter price mark-up	μ_p	$\mathcal{B}$	(0.50, 0.20)	-0.03	-0.08
MA parameter wage mark-up	μ_w	$\mathcal{B}$	(0.50, 0.20)	-0.03	-0.01
Productivity standard error	σ_a	$\mathcal{IG}$	(0.10, 2.00)	-0.02	0.03
Risk-premium standard error	σ_b	$\mathcal{IG}$	(0.10, 2.00)	0.12	-0.04
Government expenditure s.e.	σ_g	$\mathcal{IG}$	(0.10, 2.00)	-0.05	-0.13
Investment standard error	σ_i	$\mathcal{IG}$	(0.10, 2.00)	0.04	-0.04
Monetary policy standard error	σ_r	$\mathcal{IG}$	(0.10, 2.00)	-0.16	-0.10
Price mark-up standard error	σ_p	$\mathcal{IG}$	(0.10, 2.00)	-0.13	-0.13
Wage mark-up standard error	σ_w	$\mathcal{IG}$	(0.10, 2.00)	-0.08	-0.12

Notes: The table shows the mean difference of the estimated parameters ($M.D.$) and the standard deviation difference ($S.S.D.$). We also report the parameter's name (Full name), the acronym symbol (Symbol), and the prior distribution (Prior) with the mean and standard deviation (Mean, St.Dev). The number of Monte Carlo iterations is set to $I = 100$ and the number of draws is set to $D = 200000$. $\mathcal{B}$ is the Beta distribution. $\mathcal{G}$ is Gamma distribution. $\mathcal{IG}$ is the Inverse-Gamma distribution. $\mathcal{N}$ is the Normal distribution.

Table 5: *Calibrated parameters according to the literature for the open-economy model*

Full Name	Symbol	Value	Source
Discount Factor	β_i	0.9960	Del Negro and Schorfheide (2008)
Depreciation Rate of Priv. Cap.	δ_i	0.0250	Ann. Cap. Depr: 0.10
Intertemp. Elas. of Sub.	$\frac{1}{\sigma_i^c}$	0.2000	Jermann (1998)
Elast. Labour Supply	σ_i^l	4.0000	Chetty et al. (2013)
S.S. Mark-up in Goods Market	ν_i^p	1.5000	Smets and Wouters (2007)
S.S. Mark-up in Lab. Market	ν_i^w	1.5000	Smets and Wouters (2007)
Goods Market Agg. Cur.	ϑ_i^p	10.0000	Smets and Wouters (2007)
Lab. Market Agg. Cur.	ϑ_i^w	10.0000	Smets and Wouters (2007)
Bond Intermediation Cost	ϕ_i^b	0.0001	Bodenstein et al. (2011)
Depreciation Rate of Gov. Cap.	δ_i^g	0.0050	Leeper et al. (2010b)
Priv. Cap. Share in Prod.	α_i^k	0.3000	Leeper et al. (2010b)
Pub. Cap. Share in Prod.	α_i^{kg}	0.1500	Asimakopoulos et al. (2020)

Notes: The table shows the parameter's name (Full name), the acronym symbol (Symbol), the calibrated value (Value) and the source of the parameter (Source). The subscript $i = \{Canada, US\}$ indicates that the parameters have the same values for both countries.

Table 6: *Calibrated parameters according to observed data of Canada and the US*

Parameter Description	**Canada**	**US**
Unprod. Gov. Exp. / GPD	$\frac{(G_1^u)^{SS}}{(Y_1^d)^{SS}} = 0.210$	$\frac{(G_2^u)^{SS}}{(Y_2^d)^{SS}} = 0.153$
Prod. Gov. Exp. / GPD	$\frac{(G_1^p)^{SS}}{(Y_1^d)^{SS}} = 0.069$	$\frac{(G_2^p)^{SS}}{(Y_2^d)^{SS}} = 0.073$
Gov. Transfers / GDP	$\frac{(T_1)^{SS}}{(Y_1^d)^{SS}} = 0.101$	$\frac{(T_2)^{SS}}{(Y_2^d)^{SS}} = 0.117$
S.S. Capital Tax Rate	$\left(\tau_1^k\right)^{SS} = 0.341$	$\left(\tau_2^k\right)^{SS} = 0.228$
S.S. Labour Tax Rate	$\left(\tau_1^l\right)^{SS} = 0.360$	$\left(\tau_1^l\right)^{SS} = 0.259$
Weight of Cons. in Tot. Imp.	$\omega_1^{mc} = 0.223$	$\omega_2^{mc} = 0.025$
Weight of Services in Tot. Imp.	$\omega_1^{mi} = 0.944$	$\omega_2^{mi} = 0.095$
Population Size	$\zeta_1 = 0.099$	$\zeta_2 = 0.901$

Notes: The table shows the calibrated parameters according to actual data. We report the parameter values (Parameter) for Canada and the US (see online Appendix A for a detailed description of the construction of several parameters).

Table 7: *Priors and posteriors for the endogenous parameters of the open-economy model*

				Canada		**US**	
Full Name	Symbol	Prior	(Mean, St. Dev.)	Mean	[5%, 95%]	Mean	[5%, 95%]
Cons. Habit Pers.	h	$\mathcal{B}$	(0.70, 0.05)	0.32	[0.30, 0.34]	0.17	[0.15, 0.19]
Inv. Adjustment Cost	φ^{i}	$\mathcal{G}$	(4.00, 1.50)	6.22	[6.20, 6.25]	6.49	[6.47, 6.50]
Calvo Wages Prob.	ξ^w	$\mathcal{B}$	(0.60, 0.20)	0.89	[0.86, 0.92]	0.87	[0.85, 0.88]
Calvo Prices Prob.	ξ^p	$\mathcal{B}$	(0.60, 0.20)	0.82	[0.80, 0.84]	0.85	[0.83, 0.87]
Degree of Wage Ind.	ι^w	$\mathcal{B}$	(0.50, 0.15)	0.95	[0.93, 0.97]	0.67	[0.64, 0.70]
Degree of Price Ind.	ι^p	$\mathcal{B}$	(0.50, 0.15)	0.99	[0.97, 1.01]	0.19	[0.17, 0.21]
Int. Rate Smoothing	ρ	$\mathcal{B}$	(0.50, 0.20)	0.41	[0.38, 0.43]	0.62	[0.60, 0.63]
T.R. Coef. on Inf.	r^{π}	$\mathcal{G}$	(2.00, 0.25)	1.80	[1.78, 1.82]	2.20	[2.18, 2.22]
T.R. L.R. Coef. on Y	r^{y}	$\mathcal{G}$	(0.20, 0.10)	0.07	[0.04, 0.09]	0.03	[0.02, 0.04]
T.R. S.R. Coef. on Y	r^{Δ_y}	$\mathcal{G}$	(1.20, 0.05)	0.42	[0.41, 0.44]	0.49	[0.48, 0.50]
τ^l/Y Coef.	ϕ^{yl}	$\mathcal{G}$	(0.10, 0.05)	0.63	[0.61, 0.65]	0.53	[0.51, 0.54]
τ^k/Y Coef.	ϕ^{yk}	$\mathcal{G}$	(0.40, 0.20)	1.80	[1.78, 1.82]	0.66	[0.64, 0.68]
T/Y Coef.	ϕ^{yt}	$\mathcal{G}$	(0.10, 0.05)	0.24	[0.22, 0.26]	0.62	[0.60, 0.64]
τ^l/B Coef.	γ^{bl}	$\mathcal{G}$	(0.05, 0.04)	0.68	[0.66, 0.70]	0.38	[0.34, 0.41]
τ^k/B Coef.	γ^{bk}	$\mathcal{G}$	(0.30, 0.15)	1.12	[1.10, 1.15]	0.49	[0.47, 0.51]
T/B Coef.	γ^{bt}	$\mathcal{G}$	(0.50, 0.20)	1.20	[1.18, 1.22]	3.84	[3.83, 3.86]
G^p/Y Coef.	ϕ^{yg^p}	$\mathcal{G}$	(0.15, 0.05)	0.44	[0.42, 0.46]	0.59	[0.57, 0.61]
G^u/Y Coef.	ϕ^{yg^u}	$\mathcal{G}$	(0.15, 0.05)	0.14	[0.12, 0.17]	3.52	[3.50, 3.54]
G^p/B Coef.	γ^{bg^p}	$\mathcal{G}$	(0.40, 0.20)	0.04	[0.03, 0.05]	0.29	[0.26, 0.31]
G^u/B Coef.	γ^{bg^u}	$\mathcal{G}$	(0.40, 0.20)	0.44	[0.42, 0.46]	0.02	[0.01, 0.03]
Cons. / Inv. Import Sub. El.	$\frac{1+\rho^c}{\rho^c} = \frac{1+\rho^i}{\rho^i}$	$\mathcal{N}$	(1.10, 0.10)	1.07	[1.07, 1.07]	1.01	[1.01, 1.01]

Notes: The table shows the posterior means and credible intervals for the $5th$ and $95th$ percentiles. We also report the prior means and standard deviations of the endogenous parameters. Regarding the prior distributions of the endogenous parameters, $\mathcal{B}$, $\mathcal{N}$ and $\mathcal{G}$ stand for Beta, Normal and Gamma, respectively.

Table 8: *Priors and posteriors for the exogenous parameters of the open-economy model*

				Canada		**US**	
Full Name	Symbol	Prior	(Mean, St. Dev.)	Mean	[5%, 95%]	Mean	[5%, 95%]
Investment Pers.	ρ^i	$\mathcal{B}$	(0.75, 0.15)	0.94	[0.92, 0.95]	0.36	[0.34, 0.39]
Imp. Pref. Pers.	ρ^m	$\mathcal{B}$	(0.75, 0.15)	0.94	[0.92, 0.97]	0.92	[0.90, 0.94]
Wage Mark-up Pers.	ρ^w	$\mathcal{B}$	(0.75, 0.15)	0.80	[0.77, 0.82]	0.80	[0.77, 0.82]
Price Mark-up Pers.	ρ^p	$\mathcal{B}$	(0.75, 0.15)	0.86	[0.84, 0.89]	0.96	[0.94, 0.97]
Productivity Pers.	ρ^a	$\mathcal{B}$	(0.75, 0.15)	0.50	[0.49, 0.52]	0.82	[0.80, 0.85]
Prod. Gov. Exp. Pers.	ρ^{gp}	$\mathcal{B}$	(0.75, 0.15)	0.71	[0.69, 0.73]	0.91	[0.89, 0.94]
Unprod. Gov. Exp. Pers.	ρ^{gu}	$\mathcal{B}$	(0.75, 0.15)	0.77	[0.75, 0.80]	0.89	[0.88, 0.91]
Gov. Transfers Pers.	ρ^t	$\mathcal{B}$	(0.75, 0.15)	0.15	[0.13, 0.17]	0.21	[0.19, 0.23]
Capital Tax Pers.	ρ^k	$\mathcal{B}$	(0.75, 0.15)	0.53	[0.52, 0.55]	0.15	[0.13, 0.16]
Labour Income Tax Pers.	ρ^l	$\mathcal{B}$	(0.75, 0.15)	0.53	[0.50, 0.56]	0.79	[0.77, 0.80]
Monetary Policy Pers.	ρ^r	$\mathcal{B}$	(0.75, 0.15)	0.96	[0.94, 0.97]	0.97	[0.96, 0.98]
Investment St. Err.	σ^i	$\mathcal{IG}$	(0.10, 2.00)	0.18	[0.16, 0.20]	0.61	[0.58, 0.64]
Imp. Pref. St. Err.	σ^m	$\mathcal{IG}$	(0.10, 2.00)	0.05	[0.03, 0.07]	0.03	[0.02, 0.04]
Wage Mark-up St. Err.	σ^w	$\mathcal{IG}$	(0.10, 2.00)	0.07	[0.05, 0.08]	0.05	[0.03, 0.07]
Price Mark-up St. Err.	σ^p	$\mathcal{IG}$	(0.10, 2.00)	0.06	[0.04, 0.07]	0.06	[0.05, 0.08]
Productivity St. Err.	σ^a	$\mathcal{IG}$	(0.10, 2.00)	0.07	[0.06, 0.09]	0.03	[0.02, 0.05]
Prod. Gov. Exp. St. Err.	σ^{gp}	$\mathcal{IG}$	(0.10, 2.00)	0.12	[0.10, 0.14]	0.06	[0.05, 0.07]
Unprod. Gov. Exp. St. Err.	σ^{gu}	$\mathcal{IG}$	(0.10, 2.00)	0.23	[0.21, 0.25]	0.17	[0.13, 0.20]
Gov. Transfers St. Err.	σ^t	$\mathcal{IG}$	(0.10, 2.00)	0.02	[0.02, 0.03]	0.11	[0.09, 0.12]
Capital Tax St. Err.	σ^k	$\mathcal{IG}$	(0.10, 2.00)	0.05	[0.03, 0.07]	0.04	[0.03, 0.06]
Labour Income Tax St. Err.	σ^l	$\mathcal{IG}$	(0.10, 2.00)	0.03	[0.02, 0.04]	0.13	[0.11, 0.15]
Monetary Policy St. Err.	σ^r	$\mathcal{IG}$	(0.10, 2.00)	0.25	[0.23, 0.27]	1.07	[1.05, 1.09]

Notes: The table shows the posterior means and credible intervals for the $5th$ and $95th$ percentiles. We also report the prior means and standard deviations of the endogenous parameters. Regarding the prior distributions of the endogenous parameters, $\mathcal{B}$ and $\mathcal{IG}$ stand for Beta and Inverse Gamma, respectively.

Table 9: *Present-value multipliers for output and consumption under productive and unproductive government spending shocks*

Multiplier	*Impact*	*1-yr*	*3-yrs*	∞	[min, max]
Domestic Economy with Nominal Rigidities					
Productive Government Spending Present-Value Multipliers					
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	1.003	0.965	0.925	0.853	[0.838, 1.003]
$\frac{\Delta C^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	0.016	0.016	0.020	0.025	[0.014,0.025]
Unproductive Government Spending Present-Value Multipliers					
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	0.784	0.621	0.480	0.126	[0.126,0.784]
$\frac{\Delta C^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	-0.080	-0.157	-0.211	-0.353	[-0.353,-0.080]
Domestic Economy without Nominal Rigidities					
Productive Government Spending Present-Value Multipliers					
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	0.172	0.089	-0.003	-0.170	[-0.170,0.172]
$\frac{\Delta C^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	-0.083	-0.084	-0.081	-0.078	[-0.085,-0.078]
Unproductive Government Spending Present-Value Multipliers					
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	0.192	0.116	0.013	-0.407	[-0.407,0.192]
$\frac{\Delta C^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	-0.285	-0.330	-0.369	-0.534	[-0.534,-0.285]

Notes: The table shows the cumulative present-value multipliers for output and consumption based on the estimated model where $Y^d_{1,t+i}$ denotes domestic output, $C^d_{1,t+i}$ represents domestic consumption, $G^p_{1,t+i}$ is the productive government spending, $G^u_{1,t+i}$ is the unproductive government spending and Δs indicate the relative level changes of the variables with respect to their steady-state values.

Table 10: *Present-value multipliers for output and consumption under productive and unproductive government spending shocks*

	Benchmark			Larger trade openness			Smaller trade openness		
Multiplier	*Impact*	*1-yr*	*3-yrs*	*Impact*	*1-yr*	*3-yrs*	*Impact*	*1-yr*	*3-yrs*
Domestic Economy with Nominal Rigidities									
Productive Government Spending Present-Value Multipliers									
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	1.003	0.965	0.925	0.931	0.994	1.048	0.996	0.950	0.905
$\frac{\Delta C^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	0.016	0.016	0.020	0.014	0.029	0.048	0.015	0.013	0.016
Unproductive Government Spending Present-Value Multipliers									
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	0.784	0.621	0.480	0.760	0.651	0.537	0.778	0.612	0.471
$\frac{\Delta C^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	-0.080	-0.157	-0.211	-0.072	-0.128	-0.172	-0.084	-0.162	-0.216
Domestic Economy without Nominal Rigidities									
Productive Government Spending Present-Value Multipliers									
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	0.172	0.089	-0.003	0.140	0.074	-0.005	0.174	0.091	-0.003
$\frac{\Delta C^d_{1,t+i}}{\Delta G^p_{1,t+i}}$	-0.083	-0.084	-0.080	-0.091	-0.091	-0.087	-0.082	-0.084	-0.080
Unproductive Government Spending Present-Value Multipliers									
$\frac{\Delta Y^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	0.192	0.116	0.013	0.170	0.106	0.001	0.193	0.116	0.014
$\frac{\Delta C^d_{1,t+i}}{\Delta G^u_{1,t+i}}$	-0.285	-0.330	-0.370	-0.302	-0.337	-0.373	-0.284	-0.329	-0.369

Notes: The table shows the cumulative present-value multipliers for output and consumption based on the estimated model where $Y^d_{1,t+i}$ denotes domestic output, $C^d_{1,t+i}$ represents domestic consumption, $G^p_{1,t+i}$ is the productive government spending, $G^u_{1,t+i}$ is the unproductive government spending and Δs indicate the relative level changes of the variables with respect to their steady-state values.

Appendix

This Appendix describes the MitISEM and draws from Hoogerheide et al. (2012) and Baştürk et al. (2017), please refer to them for more details.

The main objective of the MitISEM is to provide an automatic and flexible method to construct a candidate density minimizing the Kullback-Leibler divergence between two densities: the target density, and the candidate density. To construct a good candidate, a mixture of Student's-t that efficiently cover the target density is estimated. The modes, scales, degrees of freedom and mixing probabilities are quickly optimized using the importance sampling (IS) weighted expectation maximization (EM) method.

Let us define $f(\theta|y)$ as the target density kernel of θ, the k-dimensional vector of interest conditioning on the data. To simplify the notation, we use $f(\theta)$. Let $g(\theta)$ be a candidate density, a mixture of H Student-t densities such that:

$$g(\theta) = g\left(\theta|\mu_h, \Sigma_h, \nu_h\right) = \sum_{h=1}^{H} \eta_h t_k\left(\theta|\mu_h, \Sigma_h, \nu_h\right), \tag{30}$$

where μ_h is a location parameter, Σ_h is a scale matrix, and ν_h is the degrees of freedom. Finally, η_h is the mixing probability of the k-dimensional Student-t components with density:

$$t_k(\theta|\mu_h, \Sigma_h, \nu_h) = \frac{\Gamma\left(\frac{\nu_h+k}{2}\right)}{\Gamma\left(\frac{\nu_h}{2}\right)(\pi\nu_h)^{k/2}} |\Sigma_h|^{-1/2} \left(1 + \frac{(\theta-\mu_h)^\top \Sigma_h^{-1}(\theta-\mu_h)}{\nu_h}\right)^{-(k+\nu_h)/2} \tag{31}$$

with $h = 1, \ldots, H$ and Σ_h is positive definite, $\eta_h \geq 0$ and $\sum_{h=1}^{H} \eta_h = 1$. The ν_h is restricted to be $\nu_h \geq 0.01$.

The MitISEM relies on the iterative construction of a mixture of Student-t as the candidate density, minimizing the Kullback-Leibler divergence between target and candidate densities. All the parameters $(\mu_h, \Sigma_h, \nu_h, \eta_h)$ are optimized jointly using an EM algorithm. This implies a large reduction in computational time and a better candidate in most applications.

The EM algorithm (Dempster et al., 1977) is a method for achieving the maximum likelihood estimates of parameters θ in models with incomplete data or latent variables. If the latent variables were observable, the computation of the maximum likelihood estimate of θ would be relatively straightforward, depending on the degree of nonlinearity of the first order conditions. The idea behind EM is to take the expectation of the objective function, in most cases the log-likelihood function, with respect to the latent variables. The expectation of the log-likelihood function is then maximized with respect to the model parameters. In many models, expectations of the latent variables depend on the model parameters θ, hence the two steps are repeated until convergence.

In the MitISEM, the EM is used to find the optimal mixture of Student-t densities for a given set of draws from a previous candidate (and their corresponding weights). We apply an IS-weighted EM algorithm to these candidate draws instead of a regular EM algorithm to posterior draws (obtained by applying the Metropolis-Hastings method to these candidate draws), since the former has three advantages. Firstly, we do not require a burn-in sample. Secondly, the use of all candidate draws (without the rejections of the MH method) helps to prevent numerical problems with estimating scale matrices of Student-t components; also, draws with relatively small, yet positive importance weights, are helpful for this purpose. Thirdly, the use of all candidate draws may lead to a better approximation.

Following Hoogerheide et al. (2012), the EM algorithm with IS weights is given by:

$$\tilde{z}_h^i \equiv \mathrm{E}\left[z_h^i \middle| \theta^i, \mu_h^{(l-1)}, \Sigma_h^{(l-1)}, \nu_h^{(l-1)}\right] = \frac{t_k\left(\theta^i | \mu_h, \Sigma_h, \nu_h\right)\ \eta_h}{\sum_{j=1}^{H} t_k\left(\theta^i | \mu_j, \Sigma_j, \nu_j\right)\ \eta_j}, \tag{32}$$

$$\widetilde{z/w}_h^i \equiv \mathrm{E}\left[\frac{z_h^i}{w_h^i} \middle| \theta^i, \mu_h^{(l-1)}, \Sigma_h^{(l-1)}, \nu_h^{(l-1)}\right] = \tilde{z}_h^i \, \frac{k+\nu_h}{\rho_h^i + \nu_h}, \tag{33}$$

$$\begin{aligned} \xi_h^i &\equiv \mathrm{E}\left[\log w_h^i \middle| \theta^i, \mu_h^{(l-1)}, \Sigma_h^{(l-1)}, \nu_h^{(l-1)}\right] = \\ &= \left[\log\left(\frac{\rho_h^i + \nu_h}{2}\right) - \psi\left(\frac{k+\nu_h}{2}\right)\right] \tilde{z}_h^i + \left[\log\left(\frac{\nu_h}{2}\right) - \psi\left(\frac{\nu_h}{2}\right)\right](1 - \tilde{z}_h^i), \end{aligned} \tag{34}$$

$$\delta_h^i \equiv \mathrm{E}\left[\frac{1}{w_h^i} \middle| \theta^i, \mu_h^{(l-1)}, \Sigma_h^{(l-1)}, \nu_h^{(l-1)}\right] = \frac{k+\nu_h}{\rho_h^i + \nu_h}\ \tilde{z}_h^i + (1 - \tilde{z}_h^i), \tag{35}$$

where i are the draws; $\rho_h^i \equiv (\theta^i - \mu_h)^\top \Sigma_h^{-1} (\theta^i - \mu_h)$; $\psi(\cdot)$ is the digamma function (the derivative of the logarithm of the gamma function $\log\Gamma(\cdot)$); $\mu_h^{(l-1)}, \Sigma_h^{(l-1)}, \nu_h^{(l-1)}, \eta_h^{(l-1)}$ are the parameters optimized in the previous $(l-1)$ EM step.

Given the expectation of the latent variables in equation (32) to (35), the parameters of each mixture component are updated using the first order conditions of the expectation of the objective function in the maximization step:

$$\mu_h^{(l)} = \left[\sum_{i=1}^{N} W^i \ \widetilde{z/w}_h^i\right]^{-1} \left[\sum_{i=1}^{N} W^i \ \widetilde{z/w}_h^i \ \theta^i\right], \tag{36}$$

$$\hat{\Sigma}_h^{(l)} = \frac{\sum_{i=1}^{N} W^i \ \widetilde{z/w}_h^i \ \left(\theta^i - \mu_h^{(l)}\right)\left(\theta^i - \mu_h^{(l)}\right)^\top}{\sum_{i=1}^{N} W^i \ \tilde{z}_h^i}, \tag{37}$$

$$\eta_h^{(l)} = \frac{\sum_{i=1}^{N} W^i \ \tilde{z}_h^i}{\sum_{i=1}^{N} W^i}. \tag{38}$$

where $W^i \equiv f(\theta^i)/g_0(\theta^i)$ are the IS weights.

Finally, $\nu_h^{(l)}$ is solved from the first order condition of ν_h:

$$-\psi(\nu_h/2) + \log(\nu_h/2) + 1 - \frac{\sum_{i=1}^{N} W^i \ \xi_h^i}{\sum_{i=1}^{N} W^i} - \frac{\sum_{i=1}^{N} W^i \ \delta_h^i}{\sum_{i=1}^{N} W^i} = 0. \tag{39}$$

The MitISEM optimizes the degrees of freedom parameter ν_h during the EM procedure to obtain a better approximation of the target density. Furthermore, the resulting values of ν_h $(h = 1, \ldots, H)$ may provide information on the shape, e.g., kurtosis of the target distribution.

MitISEM: Detailed algorithm

The MitISEM approach for obtaining an approximation to a target density:

(1) **Initialization**: Simulate draws $\theta^1, \ldots, \theta^N$ from a 'naive' candidate distribution with density g_{naive}, which is obtained as follows. Firstly, we simulate candidate draws from a Student-t distribution with density g_{mode}, where the mode is taken equal to the mode of the target density and scale matrix equal to minus the inverse Hessian of the log-target density (evaluated at the mode), and where the degrees of freedom are chosen by the user. Secondly, the mode and scale of g_{mode} are updated using the IS weighted EM algorithm, from equations (32) to equation (38). Note that g_{naive} is already a more advanced candidate than the commonly used g_{mode}; g_{mode} typically yields a substantially worse numerical efficiency than g_{naive}.

(2) **Adaptation**: Estimate the target distribution's mean and covariance matrix using IS with the draws $\theta^1, \ldots, \theta^N$ from g_{naive}. Use these estimates as the mode and scale matrix of Student-t density $g_{adaptive}$. Draw a sample $\theta^1, \ldots, \theta^N$ from this adaptive Student-t distribution with density $g_0 = g_{adaptive}$, and compute the IS weights (W^i) for this sample.

(3) **Apply the IS-weighted EM algorithm** given the latest IS weights (W^i) and the drawn sample of step (1). The output consists of the new candidate density g with optimized $\mu_h, \Sigma_h, \nu_h, \eta_h$ for $h = 1, \ldots, H$. Draw a new sample $\theta^1, \ldots, \theta^N$ from the distribution that corresponds with this proposal density and compute corresponding IS weights (W^i).

(4) **Iterate on the number of mixture components**: Given the current mixture of H components with corresponding μ_h, Σ_h, ν_h and η_h for $h = 1, \ldots, H$, take $x\%$ of the sample $\theta^1, \ldots, \theta^N$ that correspond to the highest IS weights. Construct a new mode μ_{H+1} with these draws and IS weights and scale matrix Σ_{H+1}, which are the starting values for the additional component in the mixture candidate density. This choice ensures that the new component covers a region of the parameter space in which the previous candidate mixture had relatively too little probability mass. Given the latest IS weights and the drawn sample from the current mixture of H components, apply the IS-weighted EM algorithm to optimize *each* mixture component μ_h, Σ_h, ν_h and η_h with $h = 1, \ldots, H + 1$. Draw a new sample from the mixture of $H + 1$ components and compute the corresponding IS weights.

(5) **Assess convergence of the candidate density's quality by inspecting the IS weights** and return to step (3) unless the algorithm has converged.

Cappé et al. (2008) note that there is a renewed interest in IS, due to the possibility of parallel processing implementation. Numerical efficiency in sampling methods is not only related to the efficient sample size or relative numerical efficiency, but also to the possibility of performing the simulation process in a parallel fashion. Unlike alternative methods, such as the Random Walk Metropolis or the Gibbs Sampler, IS makes use of independent draws from the candidate density, which in turn can be obtained from multicore CPUs or GPUs. See Durham and Geweke (2011) for a very novel approach. The GPU implementation of MitISEM has been explored in Baştürk et al. (2016).