Quasi-Bayesian Model Selection*

Atsushi Inoue † Mototsugu Shintani‡

Southern Methodist University Vanderbilt University

Preliminary and Incomplete

Please do not circulate

Abstract

In this paper we establish the consistency of the model selection criterion based on the quasi-marginal likelihood obtained from Laplace type estimators (LTE). We consider cases in which parameters are strongly identified and are weakly identified. Our Monte Carlo results confirm our consistency results. Our proposed procedure is applied to select among monetary macroeconomic models using US data.

*We thank Mathias Trabandt for providing the data and code.

†Department of Economics, Southern Methodist University, 3300 Dyer Street, Suite 301, Umphrey Lee Center, Dallas, TX 75275-0496. Email: ainoue@smu.edu.

‡Department of Economics Vanderbilt University VU Station B #351819 2301 Vanderbilt Place Nashville, TN 37235-1819. Email: mototsugu.shintani@vanderbilt.edu.

1 Introduction

Thanks to the development of fast computers and accessible software packages, Bayesian methods are now commonly used in the estimation of macroeconomic models. Bayesian estimators get around numerically intractable and ill-shaped likelihood functions, which maximum likelihood estimators tend to succumb to, by incorporating economically meaningful prior information. One important yet often unacknowledged fact about Bayesian methods is that likelihood functions must be correctly specified. In a recent paper, Christiano, Trabandt and Walentin (2011) propose a quasi-Bayesian approach that does not require a correctly specified likelihood function. Using the criterion function of the impulse response function matching estimator (IRFME) of Christiano, Eichenbaum and Evans (2005), they define a quasi-likelihood function and proceed with a random walk Metropolis-Hastings algorithm. Chernozhukov and Hong (2003) establish such an approach has a frequentist justification in a more general framework and call it a Laplace-type estimator (LTE). Other applications of LTEs to estimate macroeconomic models include Christiano, Eichenbaum and Trabandt (2013) and Kormilitsina and Nekipelov (2013).

When competing models are estimated by Bayesian methods, the models are often compared by their marginal likelihood. It is quite intuitive to compare models estimated by LTE using the "marginal likelihood" obtained from the LTE criterion function. In fact, Christiano, Eichenbaum and Trabandt (2013, Table 4) report the marginal likelihoods from LTE when they compare the performance of thier macroeconomic model of wage bargaining with that of a standard labor search model. In this paper, we prove that such practice is asymptotically valid in that a model with a larger value of its marginal likelihood is either correct or better approximating with probability approaching one as the sample size goes to infinity.

We consider the consistency of model selection based on marginal likelihood in two cases, (i) parameters are all strongly identified; (ii) some parameters are weakly identified; and (iii) some model parameters are set-identified. While case (i) is standard in the model selection literature (e.g., Phillips, 1996; Sin and White, 1996), cases (ii) is empirically relevant because some parameters may not be strongly identified in macroeconomic models as evidenced by Canova and Sara (2009). We consider the case

of weak identification using a device that is similar to Stock and Wright (2000) and Guerron-Quintana, Inoue and Kilian (2013). We plan to consider the case in which parameters are set identified (Moon and Schorfheide, 2009).

Our approach allows the misspecified models and is similar in spirit to the Bayesian model selection procedure considered by Schorfheide (2000). Instead of using the marginal likelihoods (or the standard posterior odds ratio) directly, Schorfheide (2000) introduces the VAR model as a reference model in the computation of loss functions so that he can compare the performance of possibly misspecified dynamic stochastic general equilibrium (DSGE) models in the Bayesian framework. The related DSGE-VAR approach of Del Negro and Schorfheide (2004, 2009) also allows DSGE models to be misspecified which results in a small weight on DSGE model obtained by maximizing the marginal likelihood of the DSGE-VAR model. An advantage of our approach is that we can directly compare the (quasi-) marginal likelihoods even if all the competing DSGE models are misspecified.

Using the frequentist framework, Corradi and Swanson (2007), Dridi, Guay and Renault (2007) and Hnatkovska, Marmer and Tang (2012), among others, propose the hypothesis testing procedures to evaluate the relative performance of possibly misspecified DSGE models. Among these studies, Hnatkovska, Marmer and Tang (2012) may be most closely related to our approach as their test statistic is based on the estimated objective function of the IRFME, namely, the estimated distance between impulse responses of the estimated DSGE model and VAR model. While our method can be interpreted as the frequentist approach, it differs from their approach as we proposes a consistent model selection procedure. Indeed, our approach is much simpler than existing testing procedures that often relies on the data dependent critical values.

The outline of this paper is as follows: Asymptotic justifications for the quasi-marginal likelihood model selection criterion are established in Section 2. Computational issues are discussed in Section 3. A small set of Monte Carlo experiments is provided in Section 4. An empirical application is illustrated in Section 5. The concluding remarks are made in Section 6. All proofs are relegated to the appendix.

2 Asymptotic Theory

Let γ denote a vector of structural impulse responses obtained from a VAR model and $f(\alpha)$ a vector of structural impulse responses implied by a DSGE model when the value of structural parameters is given by α. The impulse response function matching estimator of Christiano, Eichenbaum and Evans (2005) minimizes the criterion function

$$\hat{q}_{A,T}(\alpha) \;=\; -\frac{T}{2}(\hat{\gamma}_T - f(\alpha))'\hat{W}_T(\hat{\gamma}_T - f(\alpha))$$

where $\hat{\gamma}_T$ is a vector of structural impulse responses obtained from an estimated VAR model, and W_T is the inverse of a consistent estimator of the asymptotic variance-covariance matrix of $\hat{\gamma}_T$. Following Chernozhukov and Hong (2003), define the quasi-posterior by

$$\frac{e^{\hat{q}_{A,T}(\alpha)\pi_A(\alpha)}}{\int_A e^{\hat{q}_{A,T}(\alpha)\pi_A(\alpha)d\alpha}d\alpha}$$

where $\pi_A(\alpha)$ is the prior probability density function. This Laplace type estimator is particularly useful when the criterion function $\hat{q}_{A,T}(\alpha)$ is not numerically tractable or when extremum estimates are not reasonable.

We will propose quasi-marginal likelihoods (QML) for selecting a model and establish the consistency of the QML in two cases. Suppose that we compare DSGE models, models A and B, where models A and B are parameterized by structural parameter vectors, α and β, and imply vectors of structural impulse responses, $f(\alpha)$ and $g(\beta)$. Note that we allow for cases in which the two models are nested, overlapping and nonnested. Following Vuong (1989), we say that model A is nested by model B if there is $\tau : A \to B$ such that $f(\alpha) \;=\; g(\tau(\alpha))$ for all $\alpha \in A$. We say that the models are overlapping if they are not nested and there are $\alpha \in A$ and $\beta \in B$ such that $f(\alpha) \;=\; g(\beta)$. We say that they are nonnested if the models are neither nested or overlapping.

Define the quasi-marginal likelihood for model f by

$$m_A \;=\; \int_A e^{-\frac{T}{2}(\hat{\gamma}_T - f(\alpha))'\hat{W}_T(\hat{\gamma}_T - f(\alpha))}\pi_A(\alpha)d\alpha \;\equiv\; \int_A e^{-\frac{T}{2}\hat{q}_T(\alpha)}d\alpha,$$

and the quasi-marginal likelihood for model g by

$$m_B \;=\; \int_B e^{-\frac{T}{2}(\hat{\gamma}_T - g(\beta))'\hat{W}_T(\hat{\gamma}_T - g(\beta))}\pi_B(\beta)d\beta \;\equiv\; \int_B e^{-\frac{T}{2}\hat{q}_B(\beta)}d\beta.$$

Let

$$q_A(\alpha_0) = (\gamma_0 - f(\alpha)_0)'\Sigma^{-1}(\gamma_0 - f(\alpha_0)), \quad (1)$$

$$q_B(\beta_0) = (\gamma_0 - g(\beta_0))'\Sigma^{-1}(\gamma_0 - f(\beta_0)), \quad (2)$$

where γ_0 is a vector of population structul impulse responses, α_0 and β_0 are the (possibly pseudo) true parameter values of α and β, and Σ is the asymptotic covariance matrix of $\hat{\gamma}_T$.

We say that the quasi-marginal likelihood model selection criterion is consistent if the following property holds: $m_A > m_B$ $[m_A < m_B]$ with probability approaching one if $q_A(\alpha_0) < q_B(\beta_0)$ $[q_A(\alpha_0) > q_B(\beta_0)]$.

2.1 The Case of Strongly Identified Parameters

First, consider the case in which the parameters are strongly identified, that is, there are unique parameter values of α and β that minimize the IRFME criterion functions in population, $q_A(\alpha) = (\gamma_0 - f(\alpha))'\Sigma^{-1}(\gamma_0 - f(\alpha))$ and $q_B(\beta) = (\gamma_0 - g(\beta))'\Sigma^{-1}(\gamma_0 - g(\beta))$. Note that strong identification does not necessarily imply that the models are correctly specified because it is possible that $q_A(\alpha_0) > 0$ or $q_B(\beta_0) > 0$.

Assumption 1

(a) A and B are compact in $\Re^{p_A}$ and $\Re^{p_B}$, respectively.

(b) There exist $\alpha_0 \in \text{int}A$ and $\beta_0 \in \text{int}B$ such that for every $\epsilon > 0$

$$\inf_{\alpha \in A: \|\alpha - \alpha_0\| \geq \epsilon} q_A(\alpha) > q_A(\alpha_0),$$

$$\inf_{\beta \in B: \|\beta - \beta_0\| \geq \epsilon} q_B(\beta) > q_B(\beta_0),$$

where $q_A(\alpha)$ and $q_B(\beta)$ are defined in (1) and (2), respectively, and Σ is positive definite.

(c) $\sup_{\alpha \in A} |\hat{q}_{A,T}(\alpha) - q_A(\alpha)| = o(1)$ and $\sup_{\beta \in B} |\hat{q}_{B,T}(\beta) - q_B(\beta)| = o(1)$ almost surely, where

$$\hat{q}_{A,T}(\alpha) = (\hat{\gamma}_T - f(\alpha))'\hat{W}_T(\hat{\gamma}_T - f(\alpha)),$$

$$\hat{q}_{B,T}(\beta) = (\hat{\gamma}_T - g(\beta))'\hat{W}_T(\hat{\gamma}_T - g(\beta)).$$

(d) $f : A \to \Re^q$ and $g : B \to \Re^q$ are twice continuously differentiable in the interior of A and B, and their Jacobian matrices $\partial f(\alpha_0)/\partial \alpha'$ and $\partial g(\beta_0)/\partial \beta'$ have rank p_A and p_B, respectively.

(e) $\pi_A : A \to \Re_+$ and $\pi_B : B \to \Re_+$ are continuous in neighborhoods of α_0 and β_0 with $\pi_A(\alpha_0) > 0$ and $\pi_B(\beta_0) > 0$, respectively.

Remarks

1. We assume that the parameters are globally identified (Assumption 1b). Because Σ is positive definite (Assumption 1b) and the Jacobian matrices are of full rank (Assumption 1d), the Hessian matrices of $q_A(\alpha)$ and $q_B(\beta)$ are positive definite. Thus the parameters are also strongly identified under our assumptions.

2. Suppose that $\hat{\gamma}_T \to \gamma_0$ and $(1/T)\hat{W}_T \to \Sigma^{-1}$ almost surely. Then $\hat{q}_{A,T}(\alpha) \to q_A(\alpha)$ and $\hat{q}_{B,T}(\beta) \to q_B(\beta)$ almost surely for each $\alpha \in A$ and $\beta \in B$, respectively. Because f and g are continuously differentiable, it follows that $q_A(\alpha)$ and $q_B(\beta)$ are uniformly continuous (Assumption 1d) and that that $\hat{q}_{A,T}(\alpha)$ and $\hat{q}_{B,T}(\beta)$ are Lipschitz functions almost surely. It follows from Lemma 1 of Andrews (1992, p.246) that $\hat{q}_{A,T}(\alpha)$ and $\hat{q}_{B,T}(\beta)$ are strongly stochastic equicontinuous. By Theorem 2 of Andrews (1992, p.245), Assumption 1(c) follows from the compactness of the parameter spaces (Assumption 1a) and the strong stochastic equicontinuity.

3. Typical prior densities are continuous in macroeconomic applications, and Assumption 1(e) is likely to be satisfied.

4. Under Assumptions 1(a)–(e), the Laplace approximations to the marginal likelihoods hold:

$$\begin{aligned} m_A &\sim \left(\frac{2\pi}{T}\right)^{\frac{p_A}{2}} \exp\left(-\frac{T}{2} q_A(\alpha_0)\right) \pi_A(\alpha_0) |\nabla^2 q_A(\alpha_0)|^{-\frac{1}{2}}, \\ m_B &\sim \left(\frac{2\pi}{T}\right)^{\frac{p_B}{2}} \exp\left(-\frac{T}{2} q_B(\beta_0)\right) \pi_B(\beta_0) |\nabla^2 q_B(\beta_0)|^{-\frac{1}{2}}, \end{aligned}$$

almost surely. These approximations allow us to use the marginal likelihoods to compare $q_A(\alpha_0)$ and $q_B(\beta_0)$.

Theorem 1. Suppose that Assumption 1 holds. Then

(a) If $q_A(\alpha_0) < q_B(\beta_0)$ $[q_A(\alpha_0) > q_B(\beta_0)]$ then $m_A > m_B$ $[m_A < m_B]$ almost surely as $T \to \infty$.

(b) If $q_A(\alpha_0) = q_B(\beta_0)$ and $\dim(\alpha) < \dim(\beta)$ [$\dim(\alpha) > \dim(\beta)$], then $m_A > m_B$ [$m_A < m_B$] almost surely as $T \to \infty$.

Remarks

1. Theorem 1(a) shows that the proposed marginal likelihood model selection criterion selects the model with a larger population IRFME criterion function almost surely. This consistency result applies whether or not the models are nested or nonnested or correctly specified or misspecified. In case of nested model comparisons, the IRFME criterion function of the nesting model is expected to be lower than that of the nested model if the value of a parameter in the nested model is set to an incorrect parameter value. If the models are nonnested, it is highly likely that one of the population IRFME criterion values is lower than the other and the one with the lower population IRFME criterion will be selected. If one model is correctly specified in that its minimized population IRFME criterion function is zero while another model is misspecified in that its minimized population IRFME criterion function is positive, our model selection criterion will select the correctly specified model almost surely. Arguably, it may still make sense to minimize the IRFME criterion function even when two models are misspecified. Our model selection criterion will select the less misspecified model almost surely.

2. In the rare event in which the minimized population IRFME criterion functions take a same value, our model selection criterion will select the more parsimonious model. I think we need to provide a reference or a discuss why it makes sense. I am still not sure why it is desirable to select a model just because it has a fewer parameters. One could define a parsimony in terms of the number of state variables rather than the number of parameters, for example.

2.2 The Case of Weakly Identified Parameters

One main reason one may prefer the quasi-Bayesian method over the classical IRFME method is that parameters are not identified. See Canova and Sala (2009) for evidence of weak identificaiton in DSGE models. Following Guerron-Quintana, Inoue and Kilian (2013), we define weak identification in the minimum distance framework as follows:

Let $\alpha = [\alpha_s' \ \alpha_w']' \in A_s$ and $\beta = [\beta_s' \ \beta_w']'$ and

$$f_T(\alpha) = f_s(\alpha_s) + T^{-\frac{1}{2}} f_w(\alpha), \quad (3)$$

$$g_T(\beta) = g_s(\beta_s) + T^{-\frac{1}{2}} g_w(\beta), \quad (4)$$

where $f_s : A_s \to \Re^q$, $f_w : A \to \Re^q$, $g_s : B_s \to \Re^q$, $g_w : B_w \to \Re^q$, $A_s \subset \Re^{p_{A_s}}$ and $B_s \subset \Re^{p_{B_s}}$. α_s and β_s are strongly identified while α_w and β_w are weakly identified. We allow for the case in which all parameters are weakly identified, e.g., $p_{A_s} = 0$ and $p_{A_w} = p_A$.

Assumption 2

(a) A_s and B_s are compact in $\Re^{p_{A_s}}$ and $\Re^{p_{B_s}}$, respectively.

(b) If $p_{A_s} > 0$ or $p_{B_s} > 0$ then there exist $\alpha_{s,0} \in \text{int}A_s$ and $\beta_{s,0} \in \text{int}B_s$ such that for every $\epsilon > 0$

$$\inf_{\alpha_s \in A_s : \|\alpha_s - \alpha_{s,0}\| \geq \epsilon} q_{A_s}(\alpha_s) > q_{A_s}(\alpha_{s,0}),$$
$$\inf_{\beta_s \in B_s : \|\beta_s - \beta_{s,0}\| \geq \epsilon} q_{B_s}(\beta_s) > q_{B_s}(\beta_{s,0}),$$

where

$$q_{A_s}(\alpha_s) = (\gamma_0 - f_s(\alpha_s))' \Sigma^{-1} (\gamma_0 - f_s(\alpha_s)),$$
$$q_{B_s}(\beta_s) = (\gamma_0 - g_s(\beta_s))' \Sigma^{-1} (\gamma_0 - g_s(\beta_s)),$$

and Σ is positive definite.

(c) $\sup_{\alpha \in A} |\hat{q}_{A,T}(\alpha) - q_A(\alpha)| = o_p(1)$ and $\sup_{\beta \in B} |\hat{q}_{B,T}(\beta) - q_B(\beta)| = o_p(1)$.

(d) $f : A \to \Re^q$ and $g : B \to \Re^q$ are twice continuously differentiable in the interior of A and B, and their Jacobian matrices $\partial f_s(\alpha_{s,0})/\partial \alpha_s'$ and $\partial g_s(\beta_{s,0})/\partial \beta_s'$ have rank p_{A_s} and p_{B_s} respectively.

(e) $\pi_A : A \to \Re_+$ and $\pi_B : B \to \Re_+$ are continuous in neighborhoods of α_0 and β_0 with $\pi_A(\alpha_0) > 0$ and $\pi_B(\beta_0) > 0$, respectively.

(f) The IRFME estimators $\hat{\alpha}_T = \text{argmax}_{\alpha \in A} \hat{q}_{A,T}(\alpha)$ and $\hat{\beta}_T = \text{argmax}_{\beta \in B} \hat{q}_{B,T}(\beta)$ satisfy the first order conditions, $\nabla \hat{q}_{A,T}(\hat{\alpha}_T) = 0$ and $\nabla \hat{q}_{B,T}(\hat{\beta}_T) = 0$ with probability one.

Theorem 2. Suppose that Assumption 2 holds. Then

(a) If $q_{s,A}(\alpha_{s,0}) < q_{s,B}(\beta_{s,0})$ $[q_{s,A}(\alpha_{s,0}) > q_{s,B}(\beta_{s,0})]$ then $m_A > m_B$ $[m_A < m_B]$ with probability approaching one as $T \to \infty$.

(b) If $q_A(\alpha_0) = q_B(\beta_0)$ and $\dim(\alpha_s) < \dim(\beta_s)$ $[\dim(\alpha_s) > \dim(\beta_s)]$, then $m_A > m_B$ $[m_A < m_B]$ with probability approaching one as $T \to \infty$.

2.3 The Case of Set Identified Parameters

3 Computational Issues

We combine the random-walk Metropolis-Hastings algorithm of Chernozhukov and Hong (2003, pp.315–316) and An and Schorfheide (2007) with the method of Chib and Jeliazkov (2001) to evaluate the quasi marginal likelihood. Specifically we proceed as follows:

Step 1. Find the mode of the log of the quasi-posterior and the inverse of the negative of the Hessian of the log of the quasi-posterior at the mode, $\tilde{\alpha} = \text{argmax}_{\alpha \in A} \ln \pi_A(\alpha) + \hat{q}_{A,T}(\alpha)$ and $\tilde{\Sigma} = -[\nabla^2_\alpha (\ln \pi_A(\tilde{\alpha}) + \hat{q}_{A,T}(\tilde{\alpha})]^{-1}$.

Step 2. Draw $\alpha^{(0)}$ from $N(\tilde{\alpha}, c_0^2 \tilde{\Sigma})$. For $j = 1, ..., J$, draw α from $N(\theta^{(j-1)}, c^2 \tilde{\Sigma})$, and set $\alpha^{(j)} = \alpha$ with probability

$$r(\alpha^{(j-1)}, \alpha) = \min \left(1, \frac{\pi(\alpha) e^{\hat{q}_{A,T}(\alpha)}}{\pi(\alpha^{(j-1)}) e^{\hat{q}_{A,T}(\alpha^{(j-1)})}} \right)$$

and $\alpha^{(j)} = \alpha^{(j-1)}$ otherwise.

Step 3. For $k = 1, ..., K$, draw $\alpha^{(k)}$ from $N(\tilde{\alpha}, c^2 \tilde{\Sigma})$.

Step 4. Estimate the log of the quasi-marginal likelihood by

$$\ln \pi_A(\tilde{\alpha}) + \hat{q}_{A,T}(\tilde{\alpha}) - \ln \hat{p}_A(\tilde{\alpha}) \tag{5}$$

where

$$\hat{p}_A(\alpha) = \frac{(1/J) \sum_{j=1}^{J} r(\alpha^{(j)}, \tilde{\alpha}) \phi_{\tilde{\alpha}, c^2 \tilde{\Sigma}}(\alpha^{(j)})}{(1/K) \sum_{k=1}^{K} r(\tilde{\alpha}, \alpha^{(k)})} \tag{6}$$

and $\phi_{\mu,\Sigma}(x)$ denotes the pdf of the normal random vector with mean μ and covariance matrix Σ evaluated at x.

4 Monte Carlo Experiments

We conduct a simple Monte Carlo simulation for the purpose of investigating theoretical predictions obtain in the previous sections. Here, we employ a simple cash-in-advance (CIA) model, the canonical monetary DSGE model that has also been often used in the previous simulation studies of macroeconomic model estimation, including Nason and Cogley (1994), Schorfheide (2000) and Hnatkovska, Marmer and Tang (2012). We compute quasi-marginal likelihoods (QMLs) by matching the model's impulse responses of output growth and inflation to a technology shock and a monetary policy shock, with those obtained from artificially generated data. Since the long-run neutrality of monetary shocks is implied by the CIA model, structural shocks are identified using Blanchard and Quah's (1989) long-run restriction imposed on the bivariate VAR model of output growth and money growth. By repeating this procedure, we evaluate how often the model that generated the data is selected by QMLs.

Since this model is explained in detail by Christiano (1991) and others, here we briefly provide its basic structure and implied dynamics. The economy consists of three agents, a household, a firm and a financial intermediary. The CIA constraint implies that the household cannot purchase the consumption goods above the cash balance, $M_t - D_t + W_t H_t$, where M_t is the money stock at the beginning of period t, D_t is the (non-negative) nominal deposit at the financial intermediary, W_t is the nominal wage and H_t is hours worked. It should be noted that, in this model, the household makes the decision on the real consumption, C_t, hours worked, H_t, and deposit, D_t, after the monetary policy regarding M_{t+1} is revealed. Let E_0 denotes expectation at period 0, β denotes the discount factor and P_t denotes the aggregate price level. The representative household's maximization problem is given by

$$\max_{\{C_t, H_t, M_{t+1}, D_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^t [(1-\phi) \ln C_t + \phi \ln(1 - H_t) \right]$$

subject to $P_t C_t \leq M_t - D_t + W_t H_t$, $D_t \geq 0$ and a constraint $M_{t+1} = (M_t - D_t + W_t H_t - P_t C_t) + R_{Ht} D_t + F_t + B_t$. The last equality is the household's budget constraint which reflects the fact that the source of household's money stock in the following period includes returns on deposit with the interest rate, $R_{Ht} - 1$, nominal profits paid by the firm, F_t, and those paid by the financial intermediary, B_t.

The firm produces output Y_t using the Cobb-Douglas production technology, $Y_t = K_t^{\alpha}(Z_t N_t)^{1-\alpha}$, where K_t is the capital stock predetermined at the period t, N_t is the labor input and Z_t is labor-augmenting technology with its dynamic property introduced later. The firm must borrow $W_t N_t$ from the financial intermediary in order to pay the total wage to workers by cash. The firm repay the loan, L_t, to the financial intermediary with the interest rate $R_{Ft} - 1$ at the end of period t. Since the firm is owned by the household, the firm maximize the dividend discounted by the marginal utility of consumption. The firm solves the problem given by

$$\max_{\{F_t, K_{t+1}, N_t, L_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^{t+1} \frac{F_t}{C_{t+1} P_{t+1}} \right]$$

subject to $F_t \leq L_t + P_t \left[K_t^{\alpha}(Z_t N_t)^{1-\alpha} - K_{t+1} + (1-\delta)K_t \right] - W_t N_t - L_t R_{Ft}$ and $W_t N_t \leq L_t$. Similarly, the financial intermediary, which is also owned by the household, solves the problem given by

$$\max_{\{B_t, L_t, D_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^{t+1} \frac{B_t}{C_{t+1} P_{t+1}} \right]$$

subject to $B_t \leq D_t + R_{Ft} L_t - R_{Ht} D_t - L_t + X_t$ and $L_t \leq X_t + D_t$ where $X_t = M_{t+1} - M_t$ is the money injection by the central bank. The market clearing conditions are given by $H_t = N_t$, $P_t C_t = M_t + X_t = M_{t+1}$, $L_t = X_t + D_t = M_{t+1} - M_t + D_t$, $W_t N_t = L_t$, and $C_t + (K_{t+1} + (1-\delta)K_t) = K_t^{\alpha}(Z_t N_t)^{1-\alpha}$, and we have $R_{Ft} = R_{Ht}$ in equilibrium. From the first-order conditions of the household, we have

$$\begin{aligned} \frac{\phi}{1-\phi} \frac{C_t P_t}{1 - N_t} &= W_t \\ \frac{1}{C_t P_t} &= \beta E_t \left\{ \frac{R_t}{C_{t+1} P_{t+1}} \right\}. \end{aligned}$$

From the first-order conditions of the firm, we have

$$\begin{aligned} E_t \left\{ \frac{P_t}{C_{t+1} P_{t+1}} \right\} &= \beta E_t \left\{ \frac{P_{t+1} \alpha K_{t+1}^{\alpha-1} (Z_{t+1} N_{t+1})^{1-\alpha} + (1-\delta)}{C_{t+1} P_{t+1}} \right\} \\ R_{Ft} &= \frac{P_t (1-\alpha) K_t^{\alpha} Z_t^{1-\alpha} N_t^{-\alpha}}{W_t}. \end{aligned}$$

The money stock M_t in log is assumed follow an I(1) process given by

$$g_{M,t} = (1-\rho_M)\gamma_M + \rho_M g_{M,t-1} + \varepsilon_{M,t}$$

where $g_{M,t} = \ln(M_{t+1}/M_t)$ and $\varepsilon_{M,t} \sim N(0, \sigma_M^2)$. We also assume that the technology Z_t in log to be I(1) with its growth generated from either an AR(1) process given by

$$g_{Z,t} = (1 - \rho_Z)\gamma_Z + \rho_A g_{Z,t-1} + \varepsilon_{Z,t}$$

or an AR(2) process given by

$$g_{Z,t} = \{1 - (\rho_{Z1} + \rho_{Z2}) + \rho_{Z1}\rho_{Z2}\}\gamma_Z + (\rho_{Z1} + \rho_{Z2})g_{Z,t-1} - \rho_{Z1}\rho_{Z2}g_{Z,t-2} + \varepsilon_{Z,t}$$

where $g_{Z,t} = \ln(Z_t/Z_{t-1})$ and $\varepsilon_{Z,t} \sim N(0, \sigma_Z^2)$.

Since both M_t and Z_t contains stochastic trends, we detrend real variables by Z_t, the price level by M_t/Z_t and nominal variables by M_t. We, then, log-linearize all the conditions around the steady states and solve the dynamic system of linear difference equations which result in the theoretical impulse responses. The set of parameters to be estimated is $[\alpha,\ \beta,\ \gamma_Z,\ \exp(\gamma_M),\ \rho_Z,\ \rho_M,\ \phi,\ \delta,\ \sigma_Z,\ \sigma_M]'$ for the AR(1) technology growth specification and is $[\alpha,\ \beta,\ \gamma_Z,\ \exp(\gamma_M),\ \rho_{Z1},\ \rho_{Z2},\ \rho_M,\ \phi,\ \delta,\ \sigma_Z,\ \sigma_M]'$ for the AR(2) technology growth specification.

We consider the following two simulation designs. In the first simulation design, the artificial data of output growth, $\ln(Y_t/Y_{t-1})$, and inflation, $\ln(P_t/P_{t-1})$, are generated using the CIA model combined with the AR(1) technology growth specification using the fixed parameter values $[\alpha,\ \beta,\ \gamma_Z,\ \exp(\gamma_M),\ \rho_Z,\ \rho_M,\ \phi,\ \delta,\ \sigma_Z,\ \sigma_M]'$ =[0.33, 0.99, 0.003, 1.011, 0.7, 0.7, 0.787, 0.02, 0.005]$'$. Using the generated data for the sample size $T = 1000$, the impulse response functions are computed from the estimated bivariate VAR(4) model of $(\ln(Y_t/Y_{t-1}), \ln(P_t/P_{t-1}))'$ by imposing that the long-run impact of $\varepsilon_{M,t}$ on $\ln(Y_t/Y_{t-1})$ is zero. Both models with the AR(1) and AR(2) technology growth specification are estimated using the same data by applying the quasi-Bayesian method to match first 20 impulse responses of two variables to two shocks (so that the total number of responses are $2 \times 2 \times 20 = 80$). We label the QMLs from the AR(1) specification as m_A, and those from the AR(2) specification as m_B. Since the AR(1) model is the correct specification of technology growth, we expect that m_A takes values larger than m_B even if the AR(1) model is nested by the AR(2) model.

In the second simulation design, the data is generated from the CIA model combined with the AR(2) technology growth specification using the fixed parameter values $[\alpha,$ $\beta,\ \gamma_Z,\ \exp(\gamma_M),\ \rho_{Z1},\ \rho_{Z2},\ \rho_M,\ \phi,\ \delta,\ \sigma_Z,\ \sigma_M]'$ =[0.33, 0.99, 0.003, 1.011, 0.7, 0.7, 0.7,

0.787, 0.02, 0.005]′. The rest of the steps is the same as before. In this design, we expect that m_A takes values smaller than m_B since the AR(2) model is the correct specification for the technology growth.

The prior distribution for each model, which is given in Table 1, is similar to the one used by Schorfheide (2000) except for ρ_Z (and ρ_{Z1} and ρ_{Z2}) since he employed the iid specification for the technology growth.

The result of the simulation from 100 replications is summarized in Table 2. We report the QMLs obtained from using two alternative methods. One is based on the Laplace approximation and the other is based on the modified harmonic mean estimator. The differences of QMLs, $m_A - m_B$, are mostly positive for the simulation design 1 (AR(1) DGP) and mostly negative for the simulation design 2 (AR(2) DGP). Thus, our theoretical prediction is supported. Figure 1 also shows the empirical distribution of $m_A - m_B$ from the same experiment. The different shape of the distribution between two simulation designs reflects the fact that the nested structure has an asymmetric effect in two directions.

5 Empirical Application

In this section, we apply our procedure to a more empirically relevant medium-sized DSGE model originally developed by Christiano, Eichenbaum and Evans (2005) (hereafter CEE). In particular, we consider a modified version of the CEE model which has been estimated by Smets and Wouters (2007), Altig, Christiano, Eichenbaum and Linde (2011), Christiano, Trabandt and Walentin (2011), and Christiano, Eichenbaum and Trabandt (2013), among others. The model is one of the most commonly used macroeconomic models among practitioners that incorporates the investment adjustment cost, habit formation in consumption, sticky prices and wages, and the inflation-targeting monetary policy. In practice, this class of the model has been estimated using various methods. For example, CEE and Altig, Christiano, Eichenbaum and Linde (2011) employ the classical impulse response matching estimator, while Smets and Wouters (2007) estimate the model using a standard Bayesian method. As a third approach, Christiano, Trabandt and Walentin (2011), and Christiano, Eichenbaum and Trabandt (2013) employ the quasi-Bayesian impulse response matching estimator, or

the Laplace-type estimator.

Here, we follow Christiano, Trabandt and Walentin (2011) and Christiano, Eichenbaum and Trabandt (2013) and estimate the model using the quasi-Bayesian method. The data and estimated impulse response functions are identical to the ones used in Christiano, Trabandt and Walentin (2011).[1] They estimate a VAR(2) model of 14 variables using the US quarterly data from 1951Q1 to 2008Q4. Then, a combination of short-run and long-run restrictions is used to identify the responses to three types of shocks in the economy: (i) a monetary policy shock; (ii) a neutral technology shock; and (iii) an investment specific technology shock. The structural parameters are estimated by matching first 15 responses of selected 9 variables to 3 shocks, less 8 zero contemporaneous responses to the monetary policy shock (so that the total number of responses to match is 397).

Instead of describing the full model, we only provide equations related to three shocks. Regarding the monetary policy shock, the interest R_t is assumed follow the process given by

$$\ln(R_t/R) = \rho_R \ln(R_{t-1}/R) + (1-\rho_R)\left[r_\pi \ln(\pi_{t+1}/\pi) + r_y \ln(gdp_t/gdp)\right] + \varepsilon_{R,t}$$

where gdp_t is scaled real GDP and $\varepsilon_{R,t} \sim iid(0, \sigma_R^2)$.

In this paper, we assume that the neutral technology Z_t in log to be I(1) with its growth generated from either an iid process

$$g_{Z,t} = \gamma_Z + \varepsilon_{Z,t}$$

or an AR(1) process given by

$$g_{Z,t} = (1-\rho_Z)\gamma_Z + \rho_A g_{Z,t-1} + \varepsilon_{Z,t}$$

or an AR(2) process given by

$$g_{Z,t} = \left\{1 - (\rho_{Z1} + \rho_{Z2}) + \rho_{Z1}\rho_{Z2}\right\}\gamma_Z + (\rho_{Z1} + \rho_{Z2})g_{Z,t-1} - \rho_{Z1}\rho_{Z2}g_{Z,t-2} + \varepsilon_{Z,t}$$

where $g_{Z,t} = \ln(Z_t/Z_{t-1})$ and $\varepsilon_{Z,t} \sim iid(0, \sigma_Z^2)$.

Similarly, we assume that the investment specific technology Ψ_t in log to be I(1) with its growth generated from either an iid process

$$g_{\Psi,t} = \gamma_\Psi + \varepsilon_{\Psi,t}$$

[1] See the data appendix of their paper for the detailed explanation.

or an AR(1) process given by

$$g_{\Psi,t} = (1-\rho_\Psi)\gamma_\Psi + \rho_\Psi g_{\Psi,t-1} + \varepsilon_{\Psi,t}$$

or an AR(2) process given by

$$g_{\Psi,t} = \{1-(\rho_{\Psi 1}+\rho_{\Psi 2})+\rho_{\Psi 1}\rho_{\Psi 2}\}\gamma_\Psi + (\rho_{\Psi 1}+\rho_{\Psi 2})g_{\Psi,t-1} - \rho_{\Psi 1}\rho_{\Psi 2}g_{\Psi,t-2} + \varepsilon_{\Psi,t}$$

where $g_{\Psi,t} = \ln(\Psi_t/\Psi_{t-1})$ and $\varepsilon_{\Psi,t} \sim iid(0,\sigma_\Psi^2)$.

There are three possible specifications for the neutral technology process and three possible specifications for the investment specific technology growth process. Therefore, we have a total of 9 possible combinations regarding technology shocks. It should be noted that Christiano, Trabandt and Walentin (2011) use the iid specification for the neutral technology and the AR(1) specification for the investment specific technology. In what follows, such a combination will be denoted by iid/AR(1) specification. In contrast, Altig, Christiano, Eichenbaum and Linde (2011) and Christiano, Eichenbaum and Trabandt (2013) employ AR(1)/AR(1) specification. We compute the QML for each of 9 specifications and select the model which maximizes the QML.

The computed QMLs for 9 models are reported in Table 3. Both the Laplace approximation and the modified harmonic mean estimator suggest the AR(2)/AR(2) specification. The estimated parameters under the AR(2)/AR(2) specification are reported in Table 4, and the estimated impulse response functions are provided in Figures 2 to 4.

6 Concluding Remarks

In this paper we established the consistency of the model selection criterion based on the quasi-marginal likelihood obtained from Laplace type estimators (LTE). We considered cases in which parameters are strongly identified and are weakly identified. Our Monte Carlo results confirmed our consistency results. Our proposed procedure was also applied to select among monetary macroeconomic models using US data.

Appendix

Proof of Theorem 1.

First, we will show that the assumptions of Theorem 4.14 of Evans and Swartz (2000, p.86) hold except for a modification of their Assumption (i). Because the proofs for m_B are analogous to those for m_A, we will provide the proofs only for $\hat{q}_A(\alpha)$ to save space.

It follows from Theorem 2.1 of Newey and McFadden (1994, p.2121), their discussion on page 2122 and our Assumptions 1(a)–(c) that

$$\hat{\alpha}_T \ \to \ \alpha_0, \tag{7}$$

almost surely, and thus $\hat{\alpha}_T \ \in \ \text{int}A$ almost surely. Although Evans and Swartz (2000, p.86) assume that the parameter space is open in their Assumption (i), their proof carries through under our Assumption 1(a) of compact parameter spaces because $\hat{q}_A(\alpha)$ is continuous so that the boundary does not change the value of the integral and because $\hat{\alpha}_T$ is in the interior of A. The marginal likelihood is always well-defined because

$$\begin{aligned} m_A &= \int_A e^{-\frac{T}{2}\hat{q}_T(\alpha)}\pi_A(\alpha)d\alpha \\ &\leq \int_A \pi_A(\alpha)d\alpha \\ &= 1, \end{aligned} \tag{8}$$

where the first inequality follows from the nonnegativity of prior density $\pi_A(\alpha)$ and the positive semidefiniteness of $\hat{W}_T$. Thus, Assumption (ii) of Evans and Swartz (2000, p.86) is satisfied. It follows from (7) and our Assumptions 1(b) and 1(c) that $\hat{\alpha}_T$ uniquely maximizes $\hat{q}_A(\alpha)$ almost surely. Thus, their Assumption (iii) is satisfied. Their Assumptions (iv) and (v) are satisfied because of (7) and our Assumptions 1(d) and 1(e).

Next, we establish Laplace approximations of the marginal likelihoods. It follows from Theorem 4.14 of Evans and Swartz (2000, p.86) that

$$m_A \ = \ \left(\frac{2\pi}{T}\right)^{\frac{p_A}{2}} \exp\left(-\frac{T}{2}\hat{q}_{A,T}(\hat{\alpha}_T)\right)\pi_A(\hat{\alpha}_T)|\nabla^2\hat{q}_A(\hat{\alpha}_T)|^{-\frac{1}{2}}(1+o(1)), \tag{9}$$

It follows from (7), (9) and Assumption 1(c) that

$$m_A \;=\; \left(\frac{2\pi}{T}\right)^{\frac{p_A}{2}} \exp\left(-\frac{T}{2}q_A(\alpha_0)\right)\pi_A(\alpha_0)|\nabla^2 q_A(\alpha_0)|^{-\frac{1}{2}}(1+o(1)), \tag{10}$$

almost surely. Similarly

$$m_B \;=\; \left(\frac{2\pi}{T}\right)^{\frac{p_B}{2}} \exp\left(-\frac{T}{2}q_B(\beta_0)\right)\pi_B(\beta_0)|\nabla^2 q_B(\beta_0)|^{-\frac{1}{2}}(1+o(1)), \tag{11}$$

almost surely.

Finally, consider cases (a) and (b) in our Theorem 1. Without loss of generality suppose that $q_A(\alpha_0) < q_B(\beta_0)$ in case (a). Then the exponential terms in (10) and (11) dominate and we have $m_A > m_B$ almost surely. In case (b), the exponential rates are identical in (10) and (11) but the polynomial rates in these equations depend on the number of parameters and the model with a fewer parameters will have a larger marginal likelihood.

Proof of Theorem 2.

First, consider the case in which $\dim(\alpha_s) > 0$. Let

$$B_\epsilon(\hat{\alpha}) \;=\; \{\alpha \in A:\; \|\alpha_s - \hat{\alpha}_{s,T}\| < \epsilon_s, \|\alpha_w - \hat{\alpha}_{w,T}\| < \epsilon_w\}$$

where $\epsilon_s > 0$ and $\epsilon_w > 0$. Then

$$\begin{aligned}\left|\int_{A\setminus B_\epsilon(\hat{\alpha}_T)} \pi_A(\alpha)e^{-T\hat{q}_T(\alpha)}d\alpha\right| \;&\leq\; \int_{A\setminus B_\epsilon(\hat{\alpha}_T)} \pi_A(\alpha)d\alpha e^{-T\inf_{\alpha\in A\setminus B_\epsilon(\hat{\alpha}_T)}\hat{q}_T(\alpha)} \\ &=\; O\left(e^{-T(\hat{q}_T(\hat{\alpha}_T)+\eta)}\right)\end{aligned} \tag{12}$$

for some $\eta > 0$. By Taylor's theorem,

$$\hat{q}_T(\alpha) \;=\; \hat{q}_T(\hat{\alpha}_T) + \nabla\hat{q}_T(\hat{\alpha}_T)'(\alpha-\hat{\alpha}_T) + \frac{1}{2}(\alpha-\hat{\alpha}_T)'\nabla^2\hat{q}_T(\bar{\alpha}_T(\alpha))(\alpha-\hat{\alpha}_T) \tag{13}$$

where $\bar{\alpha}_T(\alpha)$ is a point between α and $\hat{\alpha}_T$. Thus

$$\begin{aligned}&\int_{B_\epsilon(\hat{\alpha}_T)} \pi_A(\alpha)e^{-T\hat{q}_T(\alpha)}d\alpha \\ =\;& \int_{B_\epsilon(\hat{\alpha}_T)} \pi_A(\alpha)e^{-T\hat{q}_T(\hat{\alpha}_T)+\frac{T}{2}(\alpha-\hat{\alpha}_T)'\nabla^2\hat{q}_T(\bar{\alpha}_T(\alpha))(\alpha-\hat{\alpha}_T)}d\alpha \\ =\;& \pi_A(\hat{\alpha}_T)e^{-T\hat{q}_T(\hat{\alpha}_T)}\int_{B_\epsilon(\hat{\alpha}_T)} e^{-\frac{T}{2}(\alpha-\hat{\alpha}_T)'\nabla^2\hat{q}_T(\hat{\alpha}_T)(\alpha-\hat{\alpha}_T)}d\alpha + o(\epsilon) \\ \leq\;& \sup_{\alpha_w\in A_w}\left(\pi_A([\hat{\alpha}_{s,T}'\;\alpha_w']')\right)e^{-T\hat{q}_T(\hat{\alpha}_T)}\sup_{\alpha_w\in A_w}\left\{\left|T\nabla^2\hat{q}_T([\hat{\alpha}_{s,T}'\;\alpha_w']')\right|^{-\frac{1}{2}}\right\}P(\alpha\in B_\epsilon(\hat{\alpha}_T)) + o(\epsilon),\end{aligned} \tag{14}$$

where $\alpha \sim N(\hat{\alpha}_T, [T\nabla^2 \hat{q}_T(\hat{\alpha}_T)]^{-1})$. Let $\epsilon_s + 1/(T^{1/2}\epsilon_s) = o(1)$. Then $P(\alpha \in B_\epsilon(\hat{\alpha}_T))$ is positive and is bounded away from zero. Therefore

$$\int_{B_\epsilon(\hat{\alpha}_T)} \pi_A(\alpha) e^{-T\hat{q}_T(\alpha)} d\alpha \;=\; O\left(e^{-T\hat{q}_T(\hat{\alpha}_T)} \sup_{\alpha_w \in A_w} \left\{ \left| T\nabla^2 \hat{q}_T([\hat{\alpha}'_{s,T}\ \alpha'_w]') \right|^{-\frac{1}{2}} \right\} \right). \quad (15)$$

Combining (12) and (15) yields

$$\begin{aligned} m_A &\approx e^{-T\hat{q}_T(\hat{\alpha}_T)} \sup_{\alpha_w \in A_w} \left\{ \left| T\nabla^2 \hat{q}_T([\hat{\alpha}'_{s,T}\ \alpha'_w]') \right|^{-\frac{1}{2}} \right\} && (16) \\ &\approx e^{-Tq_s A(\alpha_{s,0})} \sup_{\alpha_w \in A_w} \left\{ \left| T\nabla^2 \hat{q}_T([\alpha'_{s,0}\ \alpha'_w]') \right|^{-\frac{1}{2}} \right\} (1 + o_p(1)) && (17) \end{aligned}$$

When $\dim(\alpha_s) = 0$.

$$m_A \;=\; e^{-\frac{T}{2}\hat{\gamma}'_T \hat{W}_T \hat{\gamma}_T} (1 + o_p(1)). \quad (18)$$

The desired results follow from (17) and (18).

References

Altig, David E., Lawrence J. Christiano, Martin Eichenbaum and Jesper Lindé (2011), "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," *Review of Economic Dynamics*, 14 (2), 225–247.

An, Sungbae, and Frank Schorfheide (2007), "Bayesian Analysis of DSGE Models," *Econometric Reviews,* 26, 113–172.

Andrews, Donald W.K., (1991), "Generic Uniform Convergence," *Econometric Theory,* 8, 241–257.

Blanchard, Olivier Jean, and Danny Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673.

Canova, Fabio, and Luca Sala (2009), "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, 56, 431–449.

Chernozhukov, Victor, and Han Hong (2003), "An MCMC Approach to Classical Estimation," *Journal of Econometrics,* 293–346.

Chib, Sid, and Ivan Jeliazhov (2001), "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association,* 96, 270–281.

Christiano, Lawrence J., (1991), "Modeling the Liquidity Effect of a Money Shock," *Federal Reserve Bank of Minneapolis Quarterly Review*, 15 (1), 3–34.

Christiano, Lawrence J., Martin S. Eichenbaum and Charles Evans (2005), "Nominal Rigidities and the Dynamic Eects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

Christiano, Lawrence J., Martin S. Eichenbaum and Mathias Trabandt (2013), "Unemployment and Business Cycles," unpublished manuscript, Northwestern University and Board of Governors of the Federal Reserve System.

Christiano, Lawrence J., Mathias Trabandt and Karl Walentin (2011), "DSGE Models for Monetary Policy Analysis," in Benjamin M. Friedman, and Michael Woodford, editors: *Handbook of Monetary Economics,* Volume 3A, The Netherlands: North-Holland.

Corradi, Valentina, and Norman R. Swanson (2007), "Evaluation of Dynamic Stochastic General Equilibrium Models Based on Distributional Comparison of Simulated and Historic Data," *Journal of Econometrics*, 136 (2), 699–723.

Del Negro, Marco, and Frank Schorfheide (2004), "Priors from General Equilibrium Models for VARs," *International Economic Review*, 45(2), 643–73.

Del Negro, Marco, and Frank Schorfheide (2009), "Monetary Policy Analysis with Potentially Misspecified Models," *American Economic Review*, 99 (4), 1415–1450.

Dridi, Ramdan, Alain Guay and Eric Renault (2007), "Indirect Inference and Calibration of Dynamic Stochastic General Equilibrium Models," *Journal of Econometrics*, 136 (2), 397–430.

Evans, Michael, and Tim Swartz (2000), *Approximating Integrals via Monte Carlo and Deterministic Methods,* Oxford University Press: Oxford, UK.

Guerron-Quintana, Pablo, Atsushi Inoue and Lutz Kilian (2013), "Frequentist Inference in Weakly Identified DSGE Models," *Quantitative Economics*, forthcoming.

Kitagawa, Toru (2010), "Estimation and Inference for Set-Identified Parameters Using Posterior Lower Probability," unpublished manuscript, University College London.

Kormilitsina, Anna, and Denis Nekipelov (2013), "Consistent Variance of the Laplace Type Estimators: Application to DSGE Models," unpublished manuscript, Southern Methodist University and University of California, Berkeley.

Liao, Yuan and Wensin Jiang (2011), "Posterior Consistency of Nonparametric Conditional Moment Restricted Models," *Annals of Statistics,* 39, 3003-3031.

Meng, Xiao-Li, and Wing Hung Wong (1996), "Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration," *Statistica Sinica,* 6, 831–860.

Meng, X., and S. Schilling (2002), "Warp Bridge Sampling," *Journal of Computational and Graphical Statistics,* 11, 552–586.

Moon, Hyungsik Roger, and Frank Schorfheide (2012), "Bayesian and Frequentist Inference in Partially Identified Models," *Econometrica*, 80, 755–782.

Nason, James M., and Timothy Cogley (1994), "Testing the Implications of Long-Run Neutrality for Monetary Business Cycle Models," *Journal of Applied Econometrics*, 9 (S1), S37–S70.

Newey, Whitney K., and Daniel L. McFadden (1994), "Large Sample Estimation and Hypothesis Testing," in Robert F. Engle and Daniel L. McFadden eds., *Handbook of Econometrics,* Volume IV, 2111–2245.

Phillips, Peter C.B. (1996), "Econometric Model Determination," *Econometrica*, 64, 763–812.

Schorfheide, Frank, (2000), "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, 15 (6), 645–670.

Sin, Chor-Yiu, and Halbert White (1996), "Information Criteria for Selecting Possibly Misspecified Parametric Models," *Journal of Econometrics*, 71, 207–225.

Smets, Frank, and Rafael Wouters (2007), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

Stock, James H., and Jonathan H. Wright (2000), "GMM with Weak Identification," *Econometrica,* 68, 1055–1096.

Vuong, Quang H., (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypothesis," *Econometrica,* 57, 307–333.

Wan, Yuanyuan (2011), "An Integration-Based Approach to Moment Inequality Models," unpublished manuscript, University of Toronto.

Table 1: Prior of parameters for the CIA model

			AR(1) technology	AR(2) technology
Parameter	Distribution	[range]	Mean, std	Mean, std
α	Beta	[0,1]	0.330, 0.020	0.330, 0.020
β	Beta	[0,1]	0.990, 0.002	0.990, 0.002
γ_Z	Gaussian	$[-\infty, \infty]$	0.003, 0.003	0.003, 0.003
$\exp(\gamma_M)$	Gaussian	$[-\infty, \infty]$	1.011, 0.007	1.011, 0.007
ρ_Z	Beta	[0,1]	0.700, 0.200	-
ρ_{Z1}	Beta	[0,1]	-	0.700, 0.200
ρ_{Z2}	Uniform	[-1,1]	-	0.000, 0.577
ρ_M	Beta	[0,1]	0.700, 0.200	0.700, 0.200
ϕ	Beta	[0,1]	0.787, 0.050	0.787, 0.050
δ	Beta	[0,1]	0.010, 0.005	0.010, 0.005
σ_Z	Inv. Gamma	$[0, \infty]$	0.020, 2.000	0.020, 2.000
σ_M	Inv. Gamma	$[0, \infty]$	0.005, 2.000	0.005, 2.000

Table 2: Difference in quasi-marginal likelihoods

	DGP AR(1) technology		DGP AR(2) technology	
	Laplace approx.	MHM	Laplace approx.	MHM
mean of $m_A - m_B$	1.61	1.65	-8.19	-7.88
median of $m_A - m_B$	1.86	2.08	-6.84	-6.88
std of $m_A - m_B$	1.70	1.89	5.59	5.49
max of $m_A - m_B$	6.45	8.58	5.88	6.77
min of $m_A - m_B$	-6.23	-5.02	-30.2	-30.2
frequency of selecting true model	0.90	0.89	0.99	0.99

Notes: 100 replications. m_A is QML of the estimated model with AR(1) specification and m_B is QML of the estimated model with AR(2) specification. QMLs based on Laplace approximation (Laplace approx.) and modified harmonic mean (MHM) estimator.

Table 3: Quasi-marginal likelihoods for the CEE-type DSGE model

		Investment specific technology		
Neutral technology	Method	iid	AR(1)	AR(2)
iid	Laplace approx.	318.56	327.87	327.55
	MHM	317.75	326.73	326.36
AR(1)	Laplace approx.	324.70	334.13	333.53
	MHM	322.69	333.10	332.42
AR(2)	Laplace approx.	328.88	337.17	337.98
	MHM	327.73	334.88	335.76

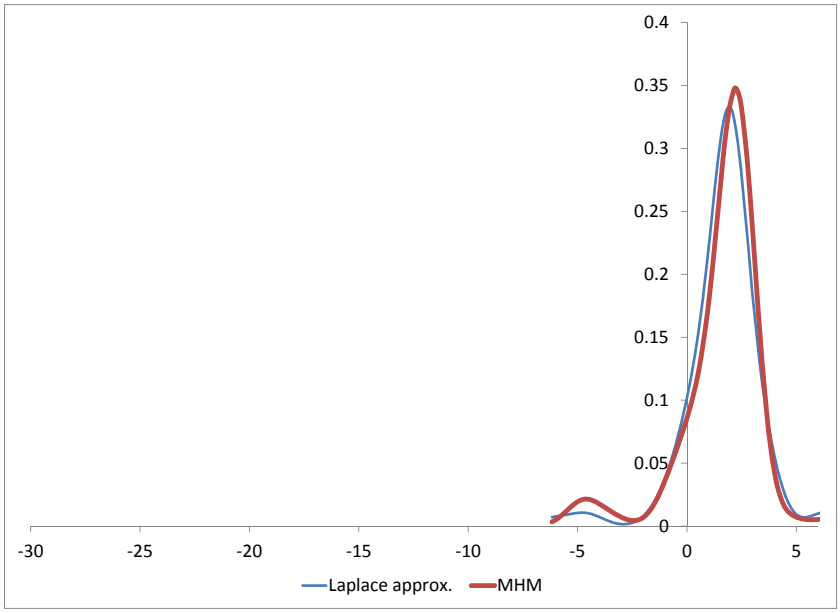
Notes: QMLs based on Laplace approximation (Laplace approx.)and modified harmonicmean (MHM) estimator.

Table 4: Prior and posteriors of parameters for the CEE-type DSGE model

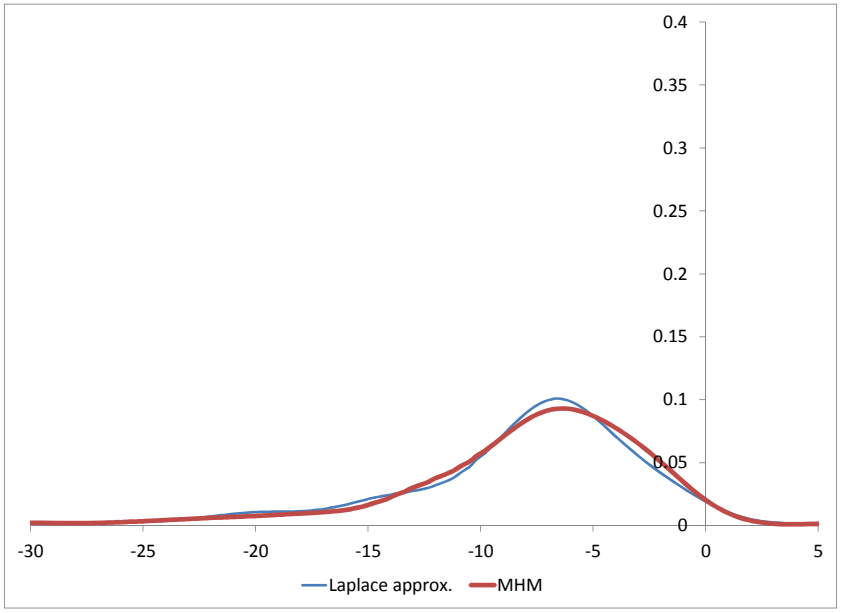
		Prior			Posterior	
Parameter		Distribution	[range]	Mean, std	Mean	[5%, 95%]
Price-setting rule						
Price stickiness	ξ_p	Beta	[0,1]	0.50, 0.15	0.54	[0.49, 0.58]
Price markup	λ_f	Gamma	[1.01, ∞]	1.20, 0.05	1.24	[1.18, 1.30]
Monetary policy rule						
Interest smoothing	ρ_R	Beta	[0,1]	0.70, 0.15	0.85	[0.83, 0.87]
Inflation coefficient	r_π	Gamma	[0, ∞]	1.70, 0.15	1.52	[1.32, 1.68]
GDP coefficient	r_y	Gamma	[0, ∞]	0.10, 0.05	0.06	[0.02, 0.11]
Preference and technology						
Consumption habit	b	Beta	[0,1]	0.50, 0.15	0.77	[0.74, 0.79]
Inv. Labor supply elast.	ϕ	Gamma	[0, ∞]	1.00, 0.50	0.86	[0.58, 1.12]
Capacity util. adj. cost	σ_a	Gamma	[0, ∞]	0.50, 0.30	0.13	[0.08, 0.21]
Investment adj. cost	S''	Gamma	[0, ∞]	8.00, 2.00	8.62	[6.29, 11.30]
Shocks						
AR(2) neutral tech.	ρ_{Z1}	Beta	[0,1]	0.20, 0.10	0.19	[0.07 , 0.35]
	ρ_{Z2}	Uniform	[-1,1]	0.00, 0.58	0.66	[0.59, 0.71]
AR(2) investment tech.	$\rho_{\Psi 1}$	Beta	[0,1]	0.20, 0.10	0.21	[0.05, 0.33]
	$\rho_{\Psi 2}$	Uniform	[-1,1]	0.00, 0.58	0.33	[0.09, 0.48]
Std neutral tech.	σ_Z	Inv. Gamma	[0, ∞]	0.20, 0.05	0.09	[0.07 , 0.10]
Std investment tech.	σ_Ψ	Inv. Gamma	[0, ∞]	0.20, 0.05	0.19	[0.15 , 0.24]
Std monetary policy	σ_R	Inv. Gamma	[0, ∞]	0.65, 0.05	0.58	[0.55 , 0.64]

Figure 1: Empirical distribution of $m_A - m_B$

(1) DGP: AR(1) technology



(2) DGP: AR(2) technology



Notes: The kernel density estimate of $m_A - m_B$ based on Laplace approximation (Laplace approx.) and modified harmonic mean (MHM) estimator.

Figure 2: Impulse responses to a monetary policy shock

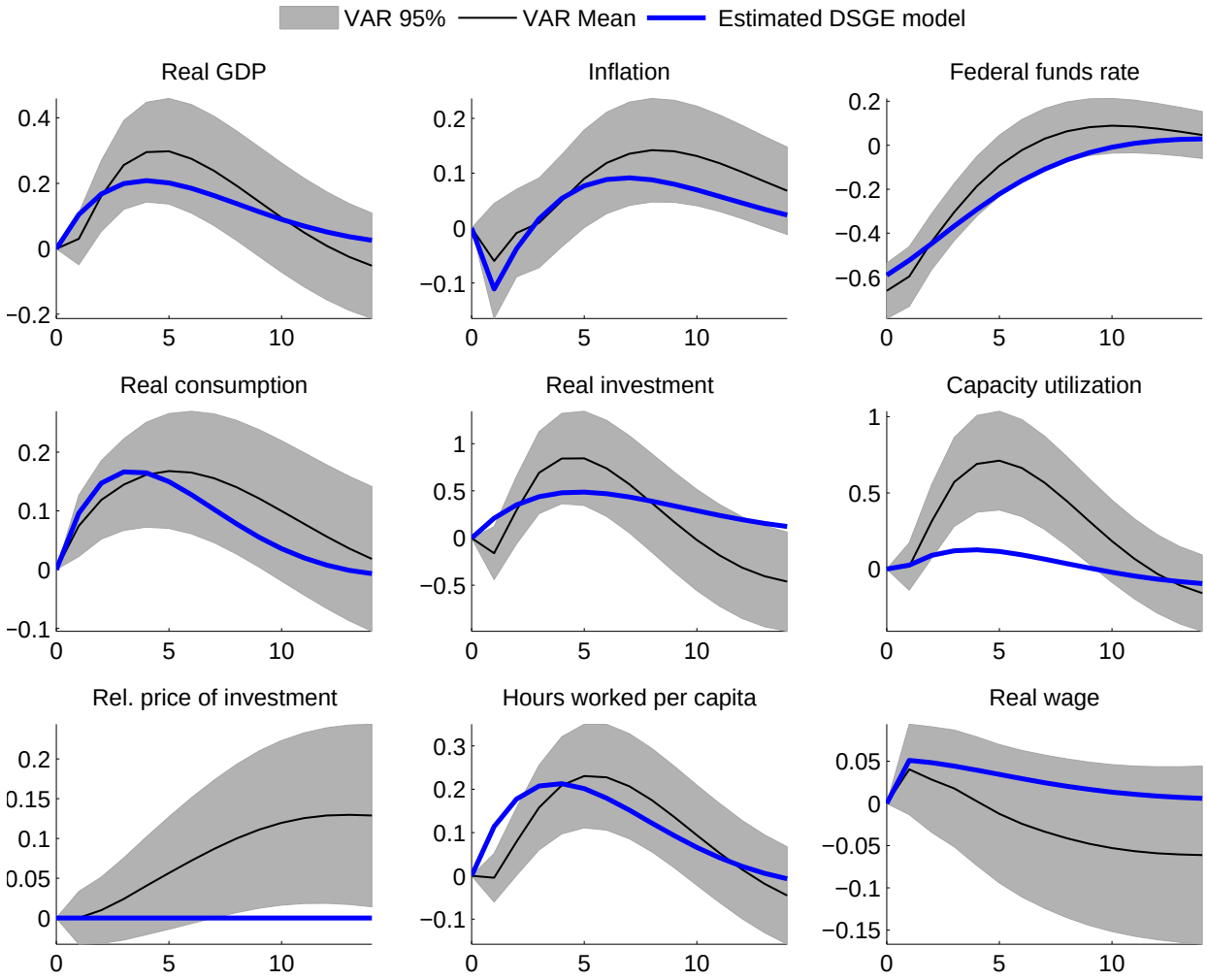


Figure 3: Impulse responses to a neutral technology shock

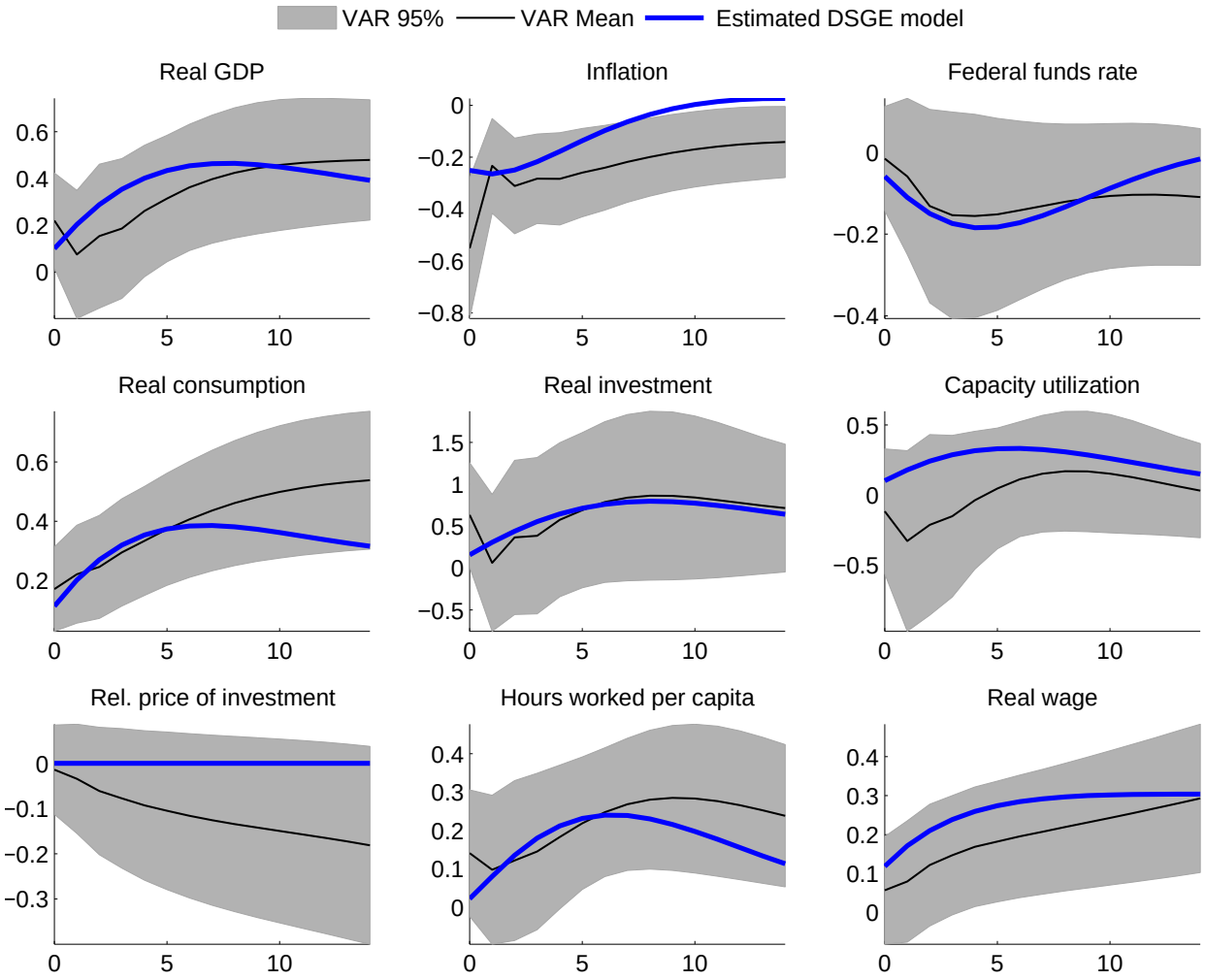


Figure 4: Impulse responses to an investment specific technology shock

