Frequency Domain Analysis of Medium Scale DSGE Models with Application to Smets and Wouters (2007)*

Denis Tkachenko† Boston University

Zhongjun Qu‡ Boston University

November 3, 2011; This version: December 16, 2012

Abstract

The paper considers parameter identification, estimation, and model diagnostics in medium scale DSGE models from a frequency domain perspective using the framework developed in Qu and Tkachenko (2012). The analysis uses Smets and Wouters (2007) as an illustrative example, motivated by the fact that it has become a workhorse model in the DSGE literature. For identification, in addition to checking parameter identifiability, we derive the non-identification curve to depict parameter values that yield observational equivalence, revealing which and how many parameters need to be fixed to achieve local identification. For estimation and inference, we contrast estimates obtained using the full spectrum with those using only the business cycle frequencies to find notably different parameter values and impulse response functions. A further comparison between the nonparametrically estimated and model implied spectra suggests that the business cycle based method delivers better estimates of the features that the model is intended to capture. Overall, the results suggest that the frequency domain based approach, in part due to its ability to handle subsets of frequencies, constitutes a flexible framework for studying medium scale DSGE models.

Keywords: Dynamic stochastic general equilibrium models, frequency domain, identification, MCMC, model diagnostics, spectrum.

*We thank participants at the 10th Annual Advances in Econometrics Conference: DSGE Models in Macroeconomics - Estimation, Evaluation, and New Developments (November 4-6, 2011), Ivan Jeliazkov, Fabio Milani and Pierre Perron for useful comments and suggestions.

†Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (tkatched@bu.edu).

‡Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (qu@bu.edu).

1 Introduction

Dynamic Stochastic General Equilibrium (DSGE) models have become a widely applied instrument for analyzing business cycles, understanding monetary policy, and for forecasting. Some medium scale DSGE models, such as that of Smets and Wouters (2007) (henceforth SW (2007)), are considered both within academia and by central banks. These models typically feature various frictions, often involving a relatively large number of equations and parameters with complex cross-equation restrictions. Although such sophistication holds promise for delivering rich and empirically relevant results, it also poses substantial challenges for identification, estimation, and inference. This paper shows how these issues can be tackled from a frequency domain perspective, using the framework recently developed by Qu and Tkachenko (2012). We use SW (2007) as the working example throughout the paper, motivated by the fact that it has become a workhorse model in the DSGE literature. The analysis of other medium scale DSGE models can be conducted in a similar manner.

The identification of DSGE models is important for both model calibration and formal statistical analysis, although the relevant literature has lagged behind relative to that concerning estimation until quite recently. Canova and Sala (2009) marks an important turning point by convincingly documenting the types of identification issues that can surface when analyzing a DSGE model. Iskrev (2010) gives sufficient conditions for the local identification of structural parameters based on the mean and a set of autocovariances. Komunjer and Ng (2011) and Qu and Tkachenko (2012) are the first to provide necessary and sufficient conditions for local identification. Qu and Tkachenko (2012) shows that taking a frequency domain perspective can deliver simple identification conditions applicable to both singular and nonsingular DSGE systems without relying on a particular (say, the minimum state) representation**.**

In this paper, we show that the methods in Qu and Tkachenko (2012) can be applied in a straightforward manner to SW (2007) to deliver informative results. We structure our identification analysis into the following steps: (1) Identification based on the second order properties. This shows whether the parameters can be identified based solely on the dynamic properties of the system. (2) Identification based on the first (i.e., the mean) and the second order properties. This reveals whether the information from the steady state restrictions can help identification. (3) Identification based on a subset of frequencies. This is motivated by the fact that DSGE models are often designed to model business cycle movements, not very long or very short term fluctuations. Upon completing the above three steps, we find that the parameters in SW (2007) are unidentified without further restrictions. (4) To obtain further insights, we derive the non-identification curves

to depict parameter values that yield observational equivalence. The curves immediately reveal which and how many parameters need to be fixed to yield local identification. Note that the results from steps (1) and (2) are in accordance with Iskrev (2010) and Komunjer and Ng (2011, the web appendix). Although these two findings are not new, the analysis is, and it also illustrates the flexibility and simplicity of taking a frequency domain approach. Issues in steps (3) and (4) have not been previously considered for medium scale DSGE models.

Next, we consider estimating SW (2007) from a frequency domain perspective using the methodology developed in Qu and Tkachenko (2012). The method has two features. First, it allows for estimation and inference using a subset of frequencies, something that is outside the scope of conventional time domain methods. This is important because DSGE models are designed for medium term economic fluctuations, not very short or long term fluctuations. Second, it is straightforward to conduct Bayesian inference and the computation involved is similar to the time domain approach. Although Qu and Tkachenko (2012) analyzed the statistical properties of this method, they did not provide an application. This paper is the first that applies the method to a medium scale DSGE model.

Specifically, we follow SW (2007) in specifying the priors and An and Schorfheide (2007) in obtaining the posterior mode and Hessian for the proposal distribution. A Random Walk Metropolis algorithm is used to obtain the posterior draws. We start with inference using the mean and the spectrum, then the full spectrum only, and finally consider inference using only business cycle frequencies. The same priors are used throughout. For the first two cases, we obtain estimates that are very similar to those of SW (2007). This reflects the close linkage between the time and frequency domain likelihood. However, for the third case, we obtain noticeably different estimates of the parameters governing the exogenous disturbances. At the same time, the parameters governing contemporaneous interactions of the observables remain similar with only a few exceptions. The impulse response functions are noticeably different. To our knowledge, this is the first time such a finding is documented in the DSGE literature.

Then, we contrast the model implied spectrum and absolute coherency with that observed in the data. The analysis is motivated by Watson's (1993) suggestion of plotting the model and data spectra as one of the most informative diagnostics. It is also related to King and Watson (1996), who compared the spectra of the three quantitative rational expectations models with that of the data. Both the business cycle and the full spectrum based estimates do a reasonable job in matching these two key features. The business cycle based estimates achieve a better fit at the

intended frequencies. However, they both underestimate the absolute coherency of the interest rate and other four variables (consumption growth, investment growth, output growth, and labor hours). The latter finding suggests a dimension along which the model can be further improved. To our knowledge, this is the first time such analysis is applied to medium scale DSGE models.

The results in the paper suggest that the frequency domain perspective affords substantial depth and flexibility in identification analysis and in estimating the parameters of the model, while remaining simple in application and comparable in terms of computational burden relative to the conventional time domain methods. In practice, we suggest to carry out both the business cycle and the full spectrum based analysis jointly. This allows us to assess to what extent the results are driven by the very low frequency contaminants, which is a hard task to tackle using a time domain framework.

The remainder of the paper is structured as follows. Section 2 includes a brief description of the SW (2007) model to make the paper self-contained. Section 3 carries out identification analysis and reports non-identification curves. Section 4 presents estimation results. Section 5 conducts model diagnostics from a frequency domain perspective. Section 6 concludes. A brief summary of the model equations not included in the text is provided in the Appendix. MATLAB code replicating the analysis is provided in an online supplement.

2 The DSGE model of SW (2007)

SW (2007) has become a workhorse model in the DSGE literature and many medium scale DSGE models consist of modifications or extensions of this model. It is an extended version of the standard New Keynesian real business cycle model, featuring a number of frictions and real rigidities. To make this paper self-contained, we subsequently briefly describe the structure of the model economy. Note that the discussion is meant to highlight the key elements in the model. For a more detailed description of the model equations, variables, and parameters, one should consult SW (2007).

The model has seven observable endogenous variables with seven exogenous shocks. In equilibrium, the model has a balanced growth path driven by deterministic labor-augmenting technological progress. We focus on the log-linearized system as in the original article. The annotated list of structural parameters can be found in Table 5.

2.1 The aggregate resource constraint

The aggregate resource constraint is given by

$$y_t = c_y c_t + i_y i_t + z_y z_t + \varepsilon_t^g.$$

Output (y_t) is composed of consumption (c_t), investment (i_t), capital utilization costs as a function of the capital utilization rate (z_t), and exogenous spending (ε_t^g). The latter is assumed to follow a first-order autoregressive model with an i.i.d. Normal error term (η_t^g), and is also affected by the fundamental productivity shock (η_t^a) as follows:

$$\varepsilon_t^g = \rho_g \varepsilon_{t-1}^g + \rho_{ga} \eta_t^a + \eta_t^g.$$

The coefficients c_y, i_y and z_y are functions of the steady state spending-output ratio (g_y), steady state output growth rate (γ), capital depreciation rate (δ), household discount factor (β), intertemporal elasticity of substitution (σ_c), fixed costs in production (ϕ_p), and share of capital in production (α) as follows: $i_y = (\gamma - 1 + \delta)k_y$, $c_y = 1 - g_y - i_y$, and $z_y = R_*^k k_y$. Here k_y is the steady state capital-output ratio, and R_*^k is the steady state rental rate of capital (see the Appendix to SW (2007)):

$$k_y = \phi_p \left(L_*/k_*\right)^{\alpha-1} = \phi_p \left[((1-\alpha)/\alpha)\left(R_*^k/w_*\right)\right]^{\alpha-1},$$

with

$$w_* = \left(\frac{\alpha^\alpha (1-\alpha)^{(1-\alpha)}}{\phi_p \left(R_*^k\right)^\alpha}\right)^{1/(1-\alpha)}$$

and

$$R_*^k = \beta^{-1}\gamma^{\sigma_c} - (1-\delta).$$

2.2 Households

Households maximize a nonseparable utility function with two arguments (consumption and labor effort) over an infinite life horizon. Consumption appears in the utility function relative to a time varying external habit variable. The dynamics of consumption follow from the consumption Euler equation

$$c_t = c_1 c_{t-1} + (1-c_1)E_t c_{t+1} + c_2(l_t - E_t l_{t+1}) - c_3(r_t - E_t \pi_{t+1}) - \varepsilon_t^b.$$

where l_t is hours worked, r_t is the nominal interest rate, and π_t is inflation. The disturbance term ε_t^b can be interpreted as a risk premium that households require to hold a one period bond. It follows the stochastic process

$$\varepsilon_t^b = \rho_b \varepsilon_{t-1}^b + \eta_t^b.$$

The relationship of the coefficients in the consumption equation to the habit persistence (λ), steady state labor market mark-up (ϕ_w), and other basic parameters highlighted above is

$$c_1 = \frac{\lambda/\gamma}{1+\lambda/\gamma}, c_2 = \frac{(\sigma_c - 1)\left(w_*^h L_*/c_*\right)}{\sigma_c\left(1+\lambda/\gamma\right)}, c_3 = \frac{1-\lambda/\gamma}{\left(1+\lambda/\gamma\right)\sigma_c},$$

where $w_*^h L_*/c_*$ is related to the steady state and is given by

$$w_*^h L_*/c_* = \frac{1}{\phi_w}\frac{1-\alpha}{\alpha}R_*^k k_y \frac{1}{c_y},$$

where R_*^k and k_y are defined as above, and $c_y = 1 - g_y - (\gamma - 1 + \delta)k_y$.

Households also choose investment given the capital adjustment cost they face. The dynamics of investment are given by

$$i_t = i_1 i_{t-1} + (1 - i_1) E_t i_{+1} + i_2 q_t + \varepsilon_t^i,$$

where ε_t^i is a disturbance to the investment specific technology process, given by

$$\varepsilon_t^i = \rho_i \varepsilon_{t-1}^i + \eta_t^i.$$

The coefficients are functions of the investment adjustment cost elasticity (φ) and other structural parameters:

$$i_1 = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}, \; i_2 = \frac{1}{\left(1+\beta\gamma^{(1-\sigma_c)}\right)\gamma^2\varphi}.$$

The corresponding arbitrage equation for the value of capital is given by

$$q_t = q_1 E_t q_{t+1} + (1 - q_1) E_t r_{t+1}^k - (r_t - \pi_{t+1}) - \frac{1}{c_3}\varepsilon_t^b \;, \tag{1}$$

with

$$q_1 = \beta\gamma^{-\sigma_c}(1-\delta) = \frac{1-\delta}{R_*^k + 1 - \delta}.$$

2.3 Final and intermediate goods market

The model has a perfectly competitive final goods market and a monopolistic competitive intermediate goods market. It features a symmetric equilibrium where all firms make identical decisions. At such an equilibrium, the aggregate production function is

$$y_t = \phi_p\left(\alpha k_t^s + (1-\alpha) l_t + \varepsilon_t^a\right),$$

where α captures the share of capital in production, and the parameter ϕ_p is one plus the fixed costs in production. Total factor productivity follows the AR(1) process

$$\varepsilon_t^a = \rho_a \varepsilon_{t-1}^a + \eta_t^a.$$

The current capital service use (k_t^s) is a function of capital installed in the previous period (k_{t-1}) and the degree of capital utilization (z_t):

$$k_t^s = k_{t-1} + z_t.$$

Furthermore, the capital utilization is a positive fraction of the rental rate of capital (r_t^k):

$$z_t = z_1 r_t^k,$$

where

$$z_1 = (1 - \psi)/\psi,$$

and ψ is the elasticity of the adjustment cost of capital utilization. The accumulation of installed capital (k_t) is given by

$$k_t = k_1 k_{t-1} + (1 - k_1)\, i_t + k_2 \varepsilon_t^i,$$

where ε_t^i is the investment specific technology process as defined before, and k_1 and k_2 are given by

$$\begin{aligned} k_1 &= \frac{1-\delta}{\gamma}, \\ k_2 &= \left(1 - \frac{1-\delta}{\gamma}\right)\left(1 + \beta\gamma^{(1-\sigma_c)}\right)\gamma^2\varphi. \end{aligned}$$

The price mark-up, defined as the difference between the average price and the nominal marginal cost, satisfies

$$\mu_t^p = \alpha\,(k_t^s - l_t) + \varepsilon_t^a - w_t,$$

where w_t is the real wage. The firms set prices according to the Calvo model, leading to the following New Keynesian Phillips curve

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 E_t \pi_{t+1} - \pi_3 \mu_t^p + \varepsilon_t^p,$$

where ε_t^p is a disturbance to the price mark-up, following the ARMA(1,1) process given by

$$\varepsilon_t^p = \rho_p \varepsilon_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p.$$

The MA(1) term is intended to pick up some of the high frequency fluctuations in prices. The Phillips curve coefficients depend on price indexation (ι_p) and stickiness (ξ_p), the curvature of the goods market Kimball aggregator (ε_p), and other structural parameters:

$$\begin{aligned}
\pi_1 &= \frac{\iota_p}{1+\beta\gamma^{(1-\sigma_c)}\iota_p}, \\
\pi_2 &= \frac{\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}\iota_p}, \\
\pi_3 &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}\iota_p}\frac{\left(1-\beta\gamma^{(1-\sigma_c)}\xi_p\right)\left(1-\xi_p\right)}{\xi_p\left(\left(\phi_p-1\right)\varepsilon_p+1\right)}.
\end{aligned}$$

Finally, cost minimization by firms implies that the rental rate of capital satisfies[1]

$$r_t^k = -\left(k_t^s - l_t\right) + w_t.$$

2.4 Labor market

Households supply their homogeneous labor to an intermediate labor union, which differentiates labor services and sets wages according to a Calvo scheme. The union then sells these services to intermediate labor packers, who in turn offer the differentiated labor package to the intermediate good producers. The wage mark-up is

$$\mu_t^w = w_t - \left(\sigma_l l_t + \frac{1}{1-\lambda}\left(c_t - \lambda c_{t-1}\right)\right),$$

where σ_l is the elasticity of labor supply with respect to real wage. Real wage w_t adjusts slowly due to the rigidity

$$w_t = w_1 w_{t-1} + \left(1-w_1\right)\left(E_t w_{t+1} + E_t\pi_{t+1}\right) - w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \varepsilon_t^w,$$

where the coefficients $w_1 - w_4$ are functions of wage indexation (ι_w) and stickiness (ξ_w) parameters, and the curvature of the labor market Kimball aggregator (ε_w):

$$\begin{aligned}
w_1 &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}, \\
w_2 &= \frac{1+\beta\gamma^{(1-\sigma_c)}\iota_w}{1+\beta\gamma^{(1-\sigma_c)}}, \\
w_3 &= \frac{\iota_w}{1+\beta\gamma^{(1-\sigma_c)}}, \\
w_4 &= \frac{1}{1+\beta\gamma^{(1-\sigma_c)}}\frac{\left(1-\beta\gamma^{(1-\sigma_c)}\xi_w\right)\left(1-\xi_w\right)}{\xi_w\left(\left(\phi_w-1\right)\varepsilon_w+1\right)}.
\end{aligned}$$

[1] In the original paper k_t instead of k_t^s shows up. In their Dynare code, SW (2007) have k_t^s.

The wage mark-up disturbance is assumed to follow an ARMA(1,1) process:

$$\varepsilon_t^w = \rho_w \varepsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w.$$

2.5 Government policies

The empirical monetary policy reaction function is

$$r_t = \rho r_{t-1} + (1-\rho)\left(r_\pi \pi_t + r_Y\left(y_t - y_t^*\right)\right) + r_{\Delta y}\left[\left(y_t - y_t^*\right) - \left(y_{t-1} - y_{t-1}^*\right)\right] + \varepsilon_t^r.$$

The monetary shock ε_t^r follows an AR(1) process:

$$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + \eta_t^r.$$

The variable y_t^* stands for a time-varying optimal output level that is the result of a flexible price-wage economy. More generally, we use superscript star to denote variables in this economy. Such an economy needs to be solved along with the sticky price-wage economy for the purposes of identification and estimation. Since the equations for the flexible price-wage economy are essentially the same as above, but with the variables μ_t^p and μ_t^w set to zero, we place them in the Appendix.

2.6 The model solution

Our analysis requires computing the spectral density matrix of the observed endogenous variables. This is straightforward to obtain using the GENSYS algorithm of Sims (2002), although other methods (e.g., Uhlig (1999)) can also be used.

The GENSYS algorithm requires representing the state variables in the following form:

$$\Gamma_0 S_t = \Gamma_1 S_{t-1} + \Psi Z_t + \Pi \zeta_t,$$

where S_t is a vector of model variables that includes the endogenous variables, the conditional expectation terms and the serially correlated exogenous shock processes, Z_t are serially uncorrelated structural disturbances, and ζ_t are expectation errors. For SW (2007) (note that the ordering of variables and parameters corresponds to our MATLAB code),

$$\begin{aligned} S_t \;=\; & [\eta_t^w, \eta_t^p, z_t^*, r_t^{k*}, k_t^{s*}, q_t^*, c_t^*, i_t^*, y_t^*, l_t^*, w_t^*, r_t^*, k_t^*, \mu_t^w, z_t, r_t^k, k_t^s, q_t, c_t, i_t, y_t, l_t, \pi_t, w_t, r_t, \\ & \varepsilon_t^a, \varepsilon_t^b, \varepsilon_t^g, \varepsilon_t^i, \varepsilon_t^r, \varepsilon_t^p, \varepsilon_t^w, k_t, E(i_{t+1}^*), E(c_{t+1}^*), E(r_{t+1}^{k*}), E(q_{t+1}^*), E(l_{t+1}^*), E(i_{t+1}), E(c_{t+1}), \\ & E(r_{t+1}^k), E(q_{t+1}), E(l_{t+1}), E(\pi_{t+1}), E(w_{t+1})]', \end{aligned}$$

where the elements 18 to 24 of S_t correspond to the observables used for identification analysis and estimation, which are (here lower cases denote log deviations from the steady state) output (y_t), consumption (c_t), investment (i_t), wage (w_t), labor hours (l_t), inflation (π_t) and the interest rate (r_t). The other elements correspond to model variables in both sticky and flexible price-wage economies, seven shock processes, and twelve expectation terms. The vector of structural shocks is given by

$$Z_t = (\eta_t^a, \eta_t^b, \eta_t^g, \eta_t^i, \eta_t^r, \eta_t^p, \eta_t^w)',$$

where, as discussed above, η_t^a is a technology shock, η_t^b is a risk premium shock, η_t^g is an exogenous spending shock, η_t^i is an investment shock, η_t^r is a monetary policy shock, η_t^p and η_t^w are price and wage mark-up shocks respectively. The elements of ζ_t are all zero except the last twelve entries that correspond to the one period ahead expectation errors of the last twelve terms of S_t. This implies that Π, which is of dimension 45×12, is an identity matrix for the last twelve rows, and zero otherwise. The coefficient matrices Γ_0, Γ_1, and Ψ are functions of the structural dynamic parameters θ, with the latter consisting of

$$\begin{aligned}\theta \;=\; & (\rho_{ga}, \mu_w, \mu_p, \alpha, \psi, \varphi, \sigma_c, \lambda, \phi_p, \iota_w, \xi_w, \iota_p, \xi_p, \sigma_l, r_\pi, r_{\Delta y}, r_y, \rho, \rho_a, \rho_b, \rho_g, \rho_i, \rho_r, \rho_p, \rho_w, \\ & \sigma_a, \sigma_b, \sigma_g, \sigma_i, \sigma_r, \sigma_p, \sigma_w, \gamma, \beta, \delta, g_y, \phi_w, \epsilon_p, \epsilon_w).\end{aligned}$$

Under conditions that ensure the existence and uniqueness of the solution (Sims (2002), p. 12), the system can be represented as

$$S_t = \Theta_1 S_{t-1} + \Theta_0 Z_t,$$

where Θ_1 and Θ_0 are functions of θ, [2] which further implies

$$S_t = (I - \Theta_1 L)^{-1} \Theta_0 Z_t. \tag{2}$$

Using the above vector moving average representation it is straightforward to obtain the representation for the observable endogenous variables. To see this, suppose that the observable Y_t, up to an unknown mean vector, is given by

$$(c_t - c_{t-1}, i_t - i_{t-1}, y_t - y_{t-1}, l_t, \pi_t, w_t - w_{t-1}, r_t). \tag{3}$$

To map this to the solution (2), we simply let $A(L)$ be a matrix of finite order lag polynomials that specifies the observables, then we compute

$$A(L) S_t = A(L)(I - \Theta_1 L)^{-1} \Theta_0 Z_t \tag{4}$$

[2] Therefore, a complete notation should be $\Theta_0(\theta)$ and $\Theta_1(\theta)$. We omit such a dependence for simplicity.

with

$$A(L)_{7\times 45} = \begin{bmatrix} \overset{(1,1)}{0} & \cdots & \overset{(1,18)}{1-L} & \overset{(1,19)}{0} & \overset{(1,20)}{0} & \overset{(1,21)}{0} & \overset{(1,22)}{0} & \overset{(1,23)}{0} & \overset{(1,24)}{0} & \cdots & \overset{(1,45)}{0} \\ \vdots & \cdots & 0 & 1-L & 0 & 0 & 0 & 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & 0 & 0 & 1-L & 0 & 0 & 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots & \vdots \\ \vdots & \cdots & 0 & 0 & 0 & 0 & 0 & 1-L & 0 & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 \end{bmatrix}.$$

Remark 1 *The vector moving average representation (4) plays a central role in our analysis. First, it enables straightforward computation of the spectrum of Y_t:*

$$f_\theta(\omega) = \frac{1}{2\pi} H(\exp(-i\omega);\theta)\Sigma(\theta)H(\exp(-i\omega);\theta)^*, \tag{5}$$

where * denotes the conjugate transpose,

$$H(L;\theta) = A(L)(I - \Theta_1 L)^{-1}\Theta_0,$$

and $\Sigma(\theta)$ is the variance covariance matrix of Z_t[3]*. Second, we can easily compute the impulse response functions and the variance decomposition. Third, the choice of $A(L)$ offers substantial flexibility as we can vary it to study estimation and inference based on different combinations of variables.*

For identification and inference based on the spectrum, there is no need to specify the steady state. However, it is also straightforward to incorporate the mean into the analysis. To see this, define an augmented parameter vector $\bar{\theta}$ that includes θ and parameters affecting only the steady state. Then, notice that for log-linearized DSGE models the observables Y_t can typically be related to the log deviations ($Y_t^d(\theta)$) and the steady states ($\mu(\bar{\theta})$) via

$$Y_t = \mu(\bar{\theta}) + Y_t^d(\theta).$$

The specification in SW (2007) corresponds to $Y_t^d(\theta)$ given by (2) and $\mu(\bar{\theta}) = (\overline{\gamma}, \overline{\gamma}, \overline{\gamma}, \bar{l}, \overline{\pi}, \overline{\gamma}, \overline{r})'$. The parameters $\overline{\gamma}, \overline{\pi}$ and $\overline{r}$ are functions of structural parameters and $\bar{l}$ is a new steady state parameter. The detailed discussion is presented in subsection 3.3 below.

[3]Note that in the code $\Sigma(\theta)$ is a 7×7 identity matrix, as the shock standard deviations are incorporated into Ψ when setting up the dynamic system.

3 Identification analysis

In this section we perform identification analysis based on the (first and) second order properties of the model. We also consider identification from a subset of frequencies (business cycle frequencies) and implement a robustness check for the results. The value of θ_0 is set to the posterior mean from the Table 1A in SW (2007):

$$\begin{aligned}\theta_0 \;=\; & (0.52, 0.88, 0.74, 0.19, 0.54, 5.48, 1.39, 0.71, 1.61, 0.59, 0.73, 0.22, 0.65, 1.92, 2.03, 0.22, 0.08,\\ & 0.81, 0.95, 0.18, 0.97, 0.71, 0.12, 0.90, 0.97, 0.45, 0.24, 0.52, 0.45, 0.24, 0.14, 0.24, 1.0043,\\ & 0.9984, 0.025, 0.18, 1.5, 10, 10).\end{aligned}$$

The above parameter values are used for illustration purposes. In practice, the same analysis can be carried out with other parameter values using the same methodology.

3.1 The identification framework

For the sake of expositional completeness, we briefly review the results in Qu and Tkachenko (2012) that are used in this section. The spectral density $f_\theta(\omega)$ plays a central role in the analysis. It can be computed using (5).

The first result concerns local identification based on the second order properties of the process. Specifically, the dynamic parameter vector θ is said to be locally identifiable from the second order properties of $\{Y_t\}$ at a point θ_0 if there exists an open neighborhood of θ_0 in which $\theta_1 \neq \theta_0$ implies $f_{\theta_1}(\omega) \neq f_{\theta_0}(\omega)$ for some $\omega \in [-\pi, \pi]$. Theorem 1 in Qu and Tkachenko (2012) establishes that a necessary and sufficient condition for local identification is that the following matrix is nonsingular:

$$G(\theta_0) = \int_{-\pi}^{\pi} \left(\frac{\partial \operatorname{vec} f_{\theta_0}(\omega)}{\partial \theta'} \right)^* \left(\frac{\partial \operatorname{vec} f_{\theta_0}(\omega)}{\partial \theta'} \right) d\omega,$$

where * stands for the conjugate transpose and the vec operator vectorizes a matrix by stacking its columns. Although the identification condition is formulated in the spectral domain, it has a time domain interpretation as well. Specifically, under some regularity condition that ensures a one-to-one mapping between the spectral density matrix and the autocovariance functions, the condition is also necessary and sufficient for local identification through the complete set of autocovariances. In practice, verifying the rank of $G(\theta)$ amounts to using an algorithm for eigenvalue decomposition and counting the number of its nonzero eigenvalues. Such a decomposition always exists because $G(\theta)$ is real, symmetric, and positive semidefinite by construction. Note that $G(\theta)$ is relatively

straightforward to compute in practice. First, the (j,k)-th element of $G(\theta)$ can be computed as

$$G_{jk}(\theta) = \int_{-\pi}^{\pi} \text{tr}\left\{ \frac{\partial f_\theta(\omega)}{\partial \theta_j} \frac{\partial f_\theta(\omega)}{\partial \theta_k} \right\} d\omega.$$

The remaining computational work is to obtain the derivatives and approximate the integral. This can be achieved using simple numerical methods. To compute the derivatives, we first divide the interval $[-\pi, \pi]$ into N subintervals to obtain $(N+1)$ frequency indices. Let ω_s denote the s-th frequency in the partition. Then one can compute $\partial f_{\theta_0}(\omega_s)/\partial \theta_j$ numerically using a simple two-point method (a refined method can also be applied to further improve precision):

$$\frac{f_{\theta_0 + \mathbf{e}_j h_j}(\omega_s) - f_{\theta_0}(\omega_s)}{h_j} \quad (j = 1, ..., N+1),$$

where $\mathbf{e}_j$ is a $q \times 1$ unit vector with the j-th element equal to 1, h_j is a step size that can be parameter dependent. In practice, to obtain the above quantity we only need to solve the DSGE model twice, once using $\theta = \theta_0$, and once with $\theta = \theta_0 + \mathbf{e}_j h_j$. After this is repeated for all parameters in θ, we can approximate the integral in $G_{jk}(\theta_0)$ using

$$\frac{2\pi}{N+1} \sum_{s=1}^{N+1} \text{tr}\left\{ \frac{\partial f_\theta(\omega_s)}{\partial \theta_j} \frac{\partial f_\theta(\omega_s)}{\partial \theta_k} \right\}.$$

Note that no simulation is needed in this process.

The identification condition can be extended to incorporate the mean (steady state properties) into the analysis. Define $\bar{\theta} = (\theta, \varkappa)'$, where the parameter vector $\varkappa$ affects only the steady state. Then, as stated in Theorem 2 of Qu and Tkachenko (2012), $\bar{\theta}$ is locally identifiable from the first and second order properties of $\{Y_t\}$ at a point $\bar{\theta}_0$ if and only if $\bar{G}(\bar{\theta}_0)$ is nonsingular, where

$$\bar{G}(\bar{\theta}) = \int_{-\pi}^{\pi} \left(\frac{\partial \,\text{vec}\, f_\theta(\omega)}{\partial \bar{\theta}'} \right)^* \left(\frac{\partial \,\text{vec}\, f_\theta(\omega)}{\partial \bar{\theta}'} \right) d\omega + \frac{\partial \mu(\bar{\theta})'}{\partial \bar{\theta}} \frac{\partial \mu(\bar{\theta})}{\partial \bar{\theta}'}.$$

In practice, the term $\mu(\bar{\theta})$ often has a simple structure and hence the derivative can be evaluated analytically, e.g., using a symbolic algebra package such as MuPAD, which is true for the SW (2007) model as will be shown later.

Qu and Tkachenko (2012) also contains two corollaries that will be useful in our analysis. First, Corollary 2 of Qu and Tkachenko (2012) provides a necessary and sufficient condition for local identification from a subset of frequencies. Specifically, let $W(\omega)$ denote an indicator function defined on $[-\pi, \pi]$ that is symmetric around zero and equal to one over a finite number of closed intervals. Also, extend the definition of $W(\omega)$ to $\omega \in [\pi, 2\pi]$ by using $W(\omega) = W(2\pi - \omega)$. Then, θ

is locally identifiable from the second order properties of $\{Y_t\}$ through the frequencies specified by $W(\omega)$ at a point θ_0 if and only if the following matrix is nonsingular:

$$G^W(\theta_0) = \left\{ \int_{-\pi}^{\pi} W(\omega) \left(\frac{\partial \operatorname{vec} f_{\theta_0}(\omega)}{\partial \theta'} \right)^* \left(\frac{\partial \operatorname{vec} f_{\theta_0}(\omega)}{\partial \theta'} \right) d\omega \right\}.$$

The result for identification of $\bar{\theta}$ is analogous. Second, Corollary 4 of Qu and Tkachenko (2012) provides a necessary and sufficient condition for conditional identification, that is, identification of a subset of parameters keeping the others fixed. Specifically, let θ^s be an s-element subset of θ, then θ^s is conditionally locally identifiable from the second order properties of $\{Y_t\}$ at a point θ_0 if and only if

$$G(\theta_0)^s = \int_{-\pi}^{\pi} \left(\frac{\partial \operatorname{vec} f_{\theta_0}(\omega)}{\partial \theta^{s\prime}} \right)^* \left(\frac{\partial \operatorname{vec} f_{\theta_0}(\omega)}{\partial \theta^{s\prime}} \right) d\omega,$$

is nonsingular. Again, the result is formulated in the same way for $\bar{\theta}^s$ using $\bar{G}(\bar{\theta}_0)^s$. It is important to note that the application of Corollary 4 does not require any additional computation once the original matrix $G(\theta_0)$ or $\bar{G}(\bar{\theta}_0)$ has been obtained. The matrices $G(\theta)^s$ or $\bar{G}(\bar{\theta}_0)^s$ for any subvector can be obtained by simply picking out the relevant elements of $G(\theta)$. Specifically, suppose we are interested in a particular k-element subvector of θ. If we number parameters inside θ, and let Φ be a set of parameter numbers of interest (i.e., if we want to vary only parameters 1,2, and 5, then $\Phi = \{1, 2, 5\}$), then the (i, j)-th element of $G(\theta)^s$ is given by

$$G(\theta)^s_{i,j} = G(\theta)_{\Phi_i, \Phi_j}, \; i = 1, 2, ..., k; \; j = 1, 2, ..., k$$

Also note that in case of Theorem 2, the same logic applies to the term

$$\frac{\partial \mu(\bar{\theta}_0)'}{\partial \bar{\theta}^s} \frac{\partial \mu(\bar{\theta}_0)}{\partial \bar{\theta}^{s\prime}}.$$

Finally, we will use the procedure Qu and Tkachenko (2012) provided to trace out non-identification curves when lack of identification is detected. The subsequent discussion focuses on θ, but the procedure works the same way for $\bar{\theta}$. Suppose $G(\theta_0)$ has only one zero eigenvalue and let $c(\theta_0)$ be the corresponding real orthonormal eigenvector. Then, $c(\theta_0)$ is unique up to multiplication by -1, and thus can be made unique by restricting its first nonzero element to be positive. Let $\delta(\theta_0)$ be an open neighborhood of θ_0. Then we can define a non-identification curve χ using the function $\theta(v)$ that solves the differential equation:

$$\begin{aligned} \frac{\partial \theta(v)}{\partial v} &= c(\theta), \\ \theta(0) &= \theta_0, \end{aligned}$$

where v is a scalar that varies in a neighborhood of 0. Then, along χ, θ is not identified at θ_0 because

$$\frac{\partial \operatorname{vec} f_{\theta(v)}(\omega)}{\partial v} = \frac{\partial \operatorname{vec} f_{\theta(v)}(\omega)}{\partial \theta(v)'} c(\theta) = 0, \ \forall \ \omega \in [-\pi, \pi].$$

Qu and Tkachenko (2012) shows that this curve is continuous and locally unique (see their Corollary 6). The curve can be evaluated numerically using any available method for solving differential equations. The simple Euler method, which amounts to recursively computing

$$\begin{aligned} \theta(v_{j+1}) &\approx \theta(v_j) + c(\theta(v_j))(v_{j+1} - v_j), \ v_{j+1} \geq v_j \geq 0, \ j = 0, 1, ... \\ \theta(v_{j-1}) &\approx \theta(v_j) + c(\theta(v_j))(v_{j-1} - v_j), \ v_{j-1} \leq v_j \leq 0, \ j = 0, -1, ..., \end{aligned} \tag{6}$$

works well in practice when setting the step size $|v_{j+1} - v_j|$ to some small number as specified below.

In cases where $G(\theta_0)$ has multiple zero eigenvalues, the following algorithm can be applied.

- **Step 1.** Apply the identification condition to verify whether all parameters in the model are locally identified. Proceed to Step 2 if lack of identification is detected.

- **Step 2.** Apply the conditional identification condition to each individual parameter. If a zero eigenvalue of $G(\theta_0)^s$ is found, then it implies that the corresponding parameter is not locally conditionally identified. Apply the procedure outlined in (6) to obtain a non-identification curve (changing only this element and fixing the value of the others at θ_0). Repeat this for all parameters to obtain a finite number of curves each being a scalar valued function of v.

- **Step 3.** Increase the number of parameters in the considered subsets of θ_0 by one at a time. Single out the subsets with the following two properties: (1) it does not include the subset detected in previous steps as a proper subset, and (2) when applying the conditional identification check, it reports only one zero eigenvalue. Repeat the procedure outlined above for all such subsets to obtain non-identification curves. Note that if the subset has k elements, then the associated curve is a k-by-1 vector valued function of v.

- **Step 4.** Continue Step 3 until all subsets are considered. Solve the model using parameter values from the curves to determine the appropriate domain for v. Truncate the curves obtained in Steps 1 to 4 to exclude parameter values contradicting economic theory or when discrepancies between $f_{\theta(v)}(\omega)$ and $f_{\theta_0}(\omega)$ are observed at some ω.

Remark 2 *The procedure above delivers two types of useful information: 1) which parameter subsets are responsible for non-identification; 2) the curve for each subset that shows in what way the parameters in this subset need to simultaneously change in order to generate observational equivalence. Considering the curves is insightful, since it allows one to go beyond traditional zero-one identification analysis and get an idea about the neighborhood of non-identification. Very small non-identified neighborhoods may not present a serious problem, but if such neighborhood spans a large portion of the parameter space including empirically relevant parameter values, then it becomes a serious issue. Such information is useful for both building a model and for estimation and inference.*

3.2 Analysis of SW (2007) based on the second order properties

To compute $G(\theta_0)$, the integral in $G(\theta_0)$ is approximated numerically by averaging over 10,000 Fourier frequencies from $-4999\pi/5000$ to $4999\pi/5000$ and multiplying by 2π. The step size for the numerical differentiation is set to $10^{-7} \times \theta_0$. The MATLAB default tolerance level $tol = \max(size(G)eps(\|G\|))$ is used to decide whether an eigenvalue is zero, where eps is the floating point precision of G. We obtain

$$\text{Rank}(G(\theta_0)) = 36.$$

Since the dimension of θ_0 is 39, this implies that θ is unidentified at θ_0. Additionally, this result suggests that a minimum of three parameters needs to be fixed to achieve identification.

Since the model is not identified, we proceed to search for the non-identified subsets of parameters. Carrying out Step 2, no such one-element subset of θ is detected. Implementing Step 3, we find two subvectors for which $G(\theta_0)^s$ has exactly one zero eigenvalue:

$$(\xi_w, \epsilon_w)$$

and

$$(\xi_p, \epsilon_p).$$

This finding is unsurprising, as the parameters in each subset play very similar roles in the model after linearization (they determine the speed of adjustment of prices and wages through the Calvo probability, or curvature of demand, respectively). They enter only jointly and thus are separately unidentifiable. Iskrev (2010) reaches the same result by applying his condition. We do not report the non-identification curves for these subsets, as they are trivial and are highlighted here for illustrative purposes.

We then exclude all three-parameter subvectors that contain either of the two non-identification sets identified above as proper subsets and continue the analysis. We find no three- or four-element non-identification subsets. However, we pinpoint one five-element subvector that yields one zero eigenvalue:

$$(\varphi, \lambda, \gamma, \beta, \delta),$$

where φ is the adjustment cost parameter, λ (denoted as h in SW (2007)) is the habit parameter, γ governs the steady state growth rate, β is the discount factor, and δ is the depreciation rate. This result is also in accordance with Iskrev (2010). After excluding all subvectors containing the non-identification sets highlighted above, we find no further sources of non-identification in this model. Therefore, our findings imply that fixing one parameter out of each of $(\varphi, \lambda, \gamma, \beta, \delta)$, (ξ_w, ϵ_w), and (ξ_p, ϵ_p) is necessary and sufficient for identification from the second order properties.

We then evaluate the non-identification curve using the Euler method with step size $h = 10^{-4} \times \theta_0$ in a small neighborhood around θ_0. The result is presented in Figure 1. It demonstrates how, for each of $\varphi, \lambda, \gamma, \beta$ and δ, the parameters have to change simultaneously in order to generate non-identification. The curve is extended using (6) in the two directions starting from θ_0(corresponding to $v = 0$ on the x-axis of the graphs), which are marked by the dotted (Direction 1) and bold (Direction 2) lines respectively. The curve is a five dimensional object. It is therefore broken down into five subplots, each corresponding to one parameter. Along Direction 1, the figure shows that increasing (φ, δ) and decreasing (λ, γ, β), while keeping the rest of the parameters fixed at their θ_0 values results in equivalent spectral densities. The values along Direction 2 can be interpreted similarly. It should be noted that β is increasing along Direction 2. Because it represents the discount factor, it should not exceed 1. Therefore, we truncate the curve before β reaches 1. This leaves us with only 472 steps in Direction 2. This, compared to 670,000 steps computed for Direction 1, is very small. Hence, in Figure 1, values corresponding to Direction 2 look like a bold dot rather than a line. Given the number of steps computed along Direction 1, we did not reach the point where natural bounds on parameters are violated, but it is clear that we would truncate it at a point where β reaches zero, λ reaches zero, or δ reaches 1, whichever happens first.

To give a further illustration of the parameter values on the curve, we report ten points taken from it at equally spaced intervals in each direction. The results are summarized in Table 1. We also compute the smallest and the second smallest eigenvalues of $G(\theta_0)^s$. The results, also reported in Table 1, show that its rank stays constant along the curve.

To verify that the points on the curve indeed result in identical spectral densities, we compute

the following three measures of discrepancies between $f_\theta(\omega)$ and $f_{\theta_0}(\omega)$ as in Qu and Tkachenko (2012):

$$\begin{aligned}
\text{Maximum absolute deviation:} \quad & \max_{\omega_j \in \Omega} | f_{\theta hl}(\omega_j) - f_{\theta_0 hl}(\omega_j)|, \\
\text{Maximum absolute deviation in relative form} \quad : \quad & \frac{\max_{\omega_j \in \Omega} | f_{\theta hl}(\omega_j) - f_{\theta_0 hl}(\omega_j)|}{| f_{\theta_0 hl}(\omega_j)|}, \\
\text{Maximum relative deviation:} \quad & \max_{\omega_j \in \Omega} \frac{| f_{\theta hl}(\omega_j) - f_{\theta_0 hl}(\omega_j)|}{| f_{\theta_0 hl}(\omega_j)|},
\end{aligned}$$

where $f_{\theta hl}(\omega)$ denotes the (h,l)-th element of the spectral density matrix with parameter θ, and Ω includes the 5,000 frequencies between 0 and π.[4] The discrepancies are summarized in Tables 2 and 3. There, the rows correspond to the parameter values reported in Table 1. The columns contain the 10 largest deviations occurring across all 49 elements of $G(\theta)^s$ in descending order. The values show that even the largest deviations are very small. Given that there are numerical errors involved in the application of GENSYS and the computation of the $G(\theta)^s$ matrix, and that the Euler method involves a cumulative approximation error of the same order as the step size (10^{-4} in our case), we can conclude that the spectrum stays the same along the curve.

3.3 Analysis of SW (2007) based on the first and the second order properties

This subsection incorporates the steady state of the model into the analysis. The measurement equations that relate the observables to the means and the log deviations are as follows (see SW (2007)):

$$\begin{aligned}
dlCONS_t &= \overline{\gamma} + c_t - c_{t-1}, \\
dlINV_t &= \overline{\gamma} + i_t - i_{t-1}, \\
dlGDP_t &= \overline{\gamma} + y_t - y_{t-1}, \\
lHOURS_t &= \bar{l} + l_t, \\
dlP_t &= \overline{\pi} + \pi_t, \\
dlWAG_t &= \overline{\gamma} + w_t - w_{t-1}, \\
FEDFUNDS_t &= \overline{r} + r_t,
\end{aligned}$$

where l and dl stand for 100 times log and log difference, respectively; $\overline{\gamma} = 100(\gamma - 1)$, $\overline{\pi} = 100(\Pi_* - 1)$, and $\overline{r} = 100(\beta^{-1}\gamma^{\sigma_c}\Pi_* - 1) = \beta^{-1}\gamma^{\sigma_c}\overline{\pi} + 100(\beta^{-1}\gamma^{\sigma_c} - 1)$. Among the means, $\overline{\gamma}$ is a

[4] There is no need to consider $\omega \in [-\pi, 0]$ because $f_\theta(\omega)$ is equal to the conjugate of $f_\theta(-\omega)$.

function of the dynamic parameter γ, $\overline{\pi}$ and $\overline{r}$ depend on the common steady parameter inflation rate Π_*, and $\bar{l}$ is a new parameter. Hence, we have

$$\overline{\theta} = (\theta, \overline{\pi}, \bar{l}).$$

There are 41 parameters in total and $\mu(\overline{\theta})$ is given by

$$\mu(\overline{\theta}) = (\overline{\gamma}, \overline{\gamma}, \overline{\gamma}, \bar{l}, \overline{\pi}, \overline{\gamma}, \overline{r})'.$$

We set $\overline{\pi}_0 = 0.78$ and $\bar{l}_0 = 0.53$ as in Table 1A in SW (2007). $\mu(\overline{\theta})$ can be differentiated analytically, e.g., using MATLAB's symbolic math toolbox.

Applying Theorem 2 yields

$$\text{Rank}(G(\overline{\theta}_0)) = 39.$$

Since now $q = 41$, $\overline{\theta}$ is unidentified at $\overline{\theta}_0$ from the first and the second order properties of the observables. Furthermore, two parameters need to be fixed to achieve identification. The sources of non-identification in this case are the two subsets we have detected in the previous subsection, namely (ξ_w, ϵ_w) and (ξ_p, ϵ_p). This result is, again, not surprising and should be expected as discussed in the previous subsection. We no longer detect the $(\varphi, \lambda, \gamma, \beta, \delta)$ subset. This is because γ determines the steady state growth rate $\overline{\gamma}$ and hence can be identified from the mean. Once γ is identified, the rest of the four parameters are uniquely determined. Iskrev (2010) reaches the same conclusion. Thus, fixing one parameter from each of (ξ_w, ϵ_w) and (ξ_p, ϵ_p) is necessary and sufficient for identification based on the mean and spectrum.

3.4 Analysis of SW (2007) using a subset of frequencies.

This subsection illustrates identification using a subset of frequencies. Without loss of generality, we focus on the business cycle frequencies (i.e., fluctuations with periods between 6 and 32 quarters as in King and Watson (1996)). We obtain

$$\text{Rank}(G^W(\theta_0)) = 36$$

and

$$\text{Rank}(\overline{G}^W(\overline{\theta}_0)) = 39,$$

which coincide with the results when all frequencies are included. All results and conclusions are the same as in the previous two subsections. This shows that for this model, business cycle frequencies have the same local identifying power at θ_0 and $\overline{\theta}_0$ as the full spectrum.

3.5 Robustness checks using non-identification curves

We first examine the sensitivity of $G(\theta_0)$ to a range of numerical differentiation steps (from $10^{-2}\times\theta_0$ to $10^{-9}\times\theta_0$) and tolerance levels (from $10^{-3}\times\theta_0$ to $10^{-10}\times\theta_0$). The results are reported in Table 4. The findings for the matrices $\bar{G}(\bar{\theta}_0), G^W(\theta_0)$ and $\overline{G}^W(\overline{\theta}_0)$ are similar and thus omitted.

Table 4 shows that different choices of step sizes and tolerance levels can affect the rank decision. Specifically, the estimated rank changes if the step size is too large or too small, and when the tolerance level is more stringent. This is quite intuitive, as when the step size is too large, the numerical differentiation will induce a substantial error, since the estimation error for the two-point method is of the same order as the step size. When the step size is too small, the numerical error from solving the model using GENSYS will be large relative to the step size, therefore the rank will also be estimated imprecisely. In this example, the step size $10^{-7}\times\theta_0$ and the MATLAB default tolerance level seem to produce good balance between precision and robustness.

The dependence of the results on the step size and the tolerance level is certainly undesirable. To address this issue, Qu and Tkachenko (2012) suggest that the non-identification curve analysis be embedded into the following two-step procedure to reduce the reliance on step size and tolerance level:

- Step 1. Compute the ranks of $G(\theta_0)$ using a wide range of step sizes and tolerance levels. Locate the outcomes with the smallest rank.
- Step 2. Derive the non-identification curves conditioning on the smallest rank reported. Compute the discrepancies in spectral densities using values on the curve to verify observational equivalence. If the discrepancies are large, proceed to the outcome with the next smallest rank and repeat the analysis. Continue until spectral densities on the curve are identical or full local identification is established.

In applications, it often suffices to compute as few as 10 points on the non-identification curve to establish whether spectral densities are identical or not, as in the latter case the deviations often become quite large only a few steps away from θ_0, so the computational burden involved is not large. Applying this procedure using the step sizes and tolerance levels in Table 4 leads to the same conclusion as stated in the previous sections. This is simply because 36 is the smallest rank in the Table (Step 1) and the discrepancies between $f_\theta(\omega)$ and $f_{\theta_0}(\omega)$ along the curves are negligible (Step 2). In summary, this example demonstrates another reason why non-identification curves can be a useful tool for identification analysis.

4 Estimation and inference

This section considers estimation of SW (2007) from a frequency domain perspective. We start with briefly summarizing the quasi-Bayesian estimation procedure proposed in Qu and Tkachenko (2012).

4.1 The basic framework

Under the assumption that the DSGE system is nonsingular, the generalized frequency domain log-likelihood function of θ based on the sample $Y_1, ..., Y_T$ is given by

$$L_T(\theta) = -\frac{1}{2}\sum_{j=1}^{T-1} W(\omega_j)\left[\log\det\left(f_\theta(\omega_j)\right) + \operatorname{tr}\left\{f_\theta^{-1}(\omega_j) I_T(\omega_j)\right\}\right],$$

where

$$\omega_j = 2\pi j/T (j = 1, 2, ...T-1)$$

denote the Fourier frequencies, $I_T(\omega_j)$ is the periodogram

$$I_T(\omega_j) = w_T(\omega_j)\, w_T^*(\omega_j)$$

with $w_T(\omega_j)$ being the discrete Fourier transform

$$w_T(\omega_j) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} Y_t \exp(-i\omega_j t),$$

and $W(\omega_j)$ is the indicator function to select frequencies for estimation and inference, as defined in the identification section.

Remark 3 *The objective function $L_T(\theta)$ allows us to estimate the dynamic parameters using the spectrum of $\{Y_t\}$ without any reference to the parameters that affect only the steady state. Also, unlike for the time domain QML, the estimates can be obtained without demeaning the data, because the values of $w_T(\omega_j)$ at the Fourier frequencies are not affected by replacing Y_t with $Y_t - \mu(\bar{\theta})$.*

The extension to the joint estimation of the dynamic and steady state parameters is straightforward. Define

$$w_{\bar{\theta},T}(0) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T}(Y_t - \mu(\bar{\theta}))$$

and

$$I_{\bar{\theta},T}(0) = w_{\bar{\theta},T}(0)\, w_{\bar{\theta},T}(0)'.$$

The log-likelihood function of $\bar{\theta}$ then takes the form

$$\bar{L}_T\left(\bar{\theta}\right) = L_T\left(\theta\right) - \frac{1}{2}\left[\log\det\left(f_\theta(0)\right) + \mathrm{tr}\left\{f_\theta^{-1}(0) I_{\bar{\theta},T}\left(0\right)\right\}\right].$$

The direct application of maximum likelihood methods to DSGE models is plagued by the problem that the parameter estimates are often at odds with economic theory, possibly due to the high dimensionality of the parameter vector and potential model misspecification. It has become common practice to introduce information not contained in the observed sample via the form of priors (see An and Schorfheide (2007) for discussion). Such an idea can be easily incorporated into the current framework.

Specifically, for the dynamic parameter only case, we consider

$$p_{_T}(\theta) = \frac{\pi(\theta)\exp\left(L_T\left(\theta\right)\right)}{\int_\Theta \pi(\theta)\exp\left(L_T\left(\theta\right)\right)d\theta},$$

where $\pi(\theta)$ can be a proper prior density or, more generally, a weight function that is strictly positive and continuous over the parameter space. The function $p_{_T}(\theta)$ is termed quasi-posterior in Chernozhukov and Hong (2003), because, while being a proper distribution density over the parameters, it is in general not a true posterior in the Bayesian sense, as $\exp\left(L_T\left(\theta\right)\right)$ is a more general criterion function than the likelihood due to the selection of the frequencies.

The estimate for θ_0 can be taken to be the quasi-posterior mean:

$$\hat{\theta}_T = \int_\Theta \theta p_{_T}(\theta) d\theta,$$

To compute the estimator, we can use Markov chain Monte Carlo (MCMC) methods, such as the Metropolis–Hastings algorithm, to draw a Markov chain

$$S = \left(\theta^{(1)}, \theta^{(2)}, ...\theta^{(B)}\right)$$

and obtain

$$\hat{\theta}_T = \frac{1}{B}\sum_{j=1}^{B}\theta^{(j)}. \tag{7}$$

Meanwhile, for a given continuously differentiable function g: $\Theta \to \mathbb{R}$ (e.g., an impulse response at a given horizon), its estimate can be obtained using

$$\hat{g}_{_T} = \frac{1}{B}\sum_{j=1}^{B} g(\theta^{(j)}). \tag{8}$$

4.2 Details on computation

In this paper, we use the Random Walk Metropolis (RWM) algorithm to generate draws from $p_T(\theta)$. It belongs to the more general class of Metropolis-Hastings algorithms, the first version of which was proposed by Metropolis et al. (1953) and later generalized by Hastings (1970). Schorfheide (2000) and Otrok (2001) were the seminal contributions in using this algorithm for Bayesian estimation of DSGE models. We use the version of the algorithm implemented in Schorfheide (2000). The main steps involved and some discussion on their practical implementation are presented below.

- Step 1. Use a numerical optimization procedure to maximize $L_T(\theta) + \log(\pi(\theta))$ to find the posterior mode, denoted $\widetilde{\theta}$.

- Step 2. Compute the inverse of the Hessian matrix evaluated at the posterior mode, denoted $\widetilde{\Sigma}$.

- Step 3. Draw a starting value $\theta^{(0)}$ from $N(\widetilde{\theta}, c^2\widetilde{\Sigma})$, where c is a scaling parameter.

- Step 4. For $s = 1, 2..., B$, draw ϑ from the proposal distribution $N(\theta^{(s-1)}, c^2\widetilde{\Sigma})$. Accept the draw $(\theta^{(s)} = \vartheta)$ with probability

$$\min\{1, \alpha(\theta^{(s-1)}, \vartheta \mid Y)\}$$

with

$$\alpha(\theta^{(s-1)}, \vartheta \mid Y) = \frac{\exp(L_T(\vartheta))\pi(\vartheta)}{\exp\left(L_T(\theta^{(s-1)})\right)\pi(\theta^{(s-1)})}.$$

If it is rejected, then set $\theta^{(s)} = \theta^{(s-1)}$.

- Step 5. Compute the posterior mean as in (7) and (8).

In Step 1, one of the practical problems is that the solution for the DSGE model may be non-unique or may not exist at some parameter values. To circumvent these issues, we use the csminwel optimization routine written by Chris Sims (see Leeper and Sims (1994)), which randomly perturbs the search direction if it reaches a cliff caused by indeterminacy or nonexistence. Regarding the prior, we use the same $\pi(\theta)$ as in Table 1A in SW (2007).

In Step 2, the Hessian matrix, computed assuming Normality, has its (j, l)-th element given by

$$\frac{1}{4\pi}\int_{-\pi}^{\pi} W(\omega)\,\mathrm{tr}\left[f_{\widetilde{\theta}}(\omega)\frac{\partial f_{\widetilde{\theta}}^{-1}(\omega)}{\partial\theta_j}f_{\widetilde{\theta}}(\omega)\frac{\partial f_{\widetilde{\theta}}^{-1}(\omega)}{\partial\theta_l}\right]d\omega,$$

which can be estimated by replacing the integral with an average over the Fourier frequencies.

In Step 4, the choice of the scaling parameter c is determined by calibrating the acceptance rate of the Markov chain. Roberts et al. (1997) suggested tuning the proposal distributions so that the acceptance rate is close to 25% for models of dimension higher than two under the assumption that both the target and the proposal distribution are Normal. Since this assumption is not satisfied in our case, we follow the literature by running several independent chains and choosing c such that the acceptance rates fall between 20% and 40%. In our experience, for a given c, computing the acceptance rate of a chain of 1,000-5,000 draws gives a good idea about what to expect from a much longer chain.

The parameter draws from the proposal distribution $N(\theta^{(s-1)}, c^2\widetilde{\Sigma})$ may yield indeterminacy or nonexistence of the DSGE solution, or fall outside of the specified bounds (our bounds are as in the Dynare code of SW (2007)). In such cases, we set $L_T(\theta) + \log(\pi(\theta))$ to a very large negative number (-10^{10}) so that such draws are always rejected.

Below, we first estimate $\bar{\theta}$ using both the mean and the full spectrum of observables, as this closely mirrors the analysis of SW (2007) conducted from a time domain perspective. In order to enhance comparability of results, we fix five parameters as in SW (2007), at the following values

$$\epsilon_p = \epsilon_w = 10, \delta = 0.025, g_y = 0.18, \lambda_w = 1.5.$$

4.3 Estimation based on the mean and the full spectrum

We use the dataset from SW (2007) and consider the same sample period as in their Dynare code, namely 1965.1-2004.4. The prior distribution is the same as in SW (2007) and is presented in Table 5. A sample of 250,000 draws from the posterior distribution is generated, and the first 50,000 are discarded. The scaling factor c is set to $\sqrt{0.15}$, which yields a rejection rate of 24%.[5] It should also be noted that the spectral density at frequency zero is singular, because the observables contain first differences of stationary variables. Computationally, we deal with this problem by using the generalized inverse to calculate $f_\theta^{-1}(0)$ and the product of nonzero eigenvalues of $f_\theta(0)$ to obtain $\det(f_\theta(0))$.

Table 6 presents the estimates and their 90% probability intervals. The results obtained in SW (2007) are also included for ease of comparison. Overall, the parameter estimates are very similar to their counterparts in SW (2007). In particular, the posterior means and modes are close and the

[5] Here and below we used several scaling factors yielding the acceptance rates between 20% and 40%, and found that the results are not sensitive to these changes.

90% probability intervals overlap for 39 out of the 41 parameters. For the latter, the two exceptions are that our estimate of the technology shock persistence (ρ_a) is higher (0.98 compared to 0.95 in SW (2007)), while the estimated persistence parameter of the exogenous spending shock (ρ_g) is lower (0.92 versus 0.97). For these two parameters the corresponding 90% probability intervals are disjoint. We can also single out a somewhat higher estimate of the elasticity of consumption σ_c (1.61 compared to 1.38), although there is still overlap in the 90% intervals, and a lower estimate of the trend growth rate ($\overline{\gamma}$) of 0.31 versus 0.43 in SW (2007).

4.4 Estimation based on the full spectrum

We now consider the estimation of θ based on the full spectrum. We use the same data set, prior, and MCMC algorithm, except we use $c = \sqrt{0.15}$, which produces an acceptance rate of 25%. The results are reported in Table 7.

Overall, the parameter estimates are very similar to those based on the mean and the full spectrum. The estimated trend growth rate is back in line with the results of SW (2007), but the estimated mean discount rate goes up from 0.84% to 1.04% on an annual basis. The rest of the estimates obtained using the full spectrum are virtually the same as those in Table 6. Consequently, overall the results using the full spectrum are very close to those obtained by SW (2007) using time domain methods.

4.5 Estimation using business cycle frequencies

DSGE models are constructed to explain business cycle movements. Schorfheide (2011) emphasized that "many time series exhibit low frequency behavior that is difficult, if not impossible, to reconcile with the model being estimated. This low frequency misspecification contaminates the estimation of shocks and thereby inference about the sources of business cycle". Therefore, it is instructive to examine how the parameters change when they are estimated using only business cycles frequencies. Our procedure allows for such an investigation. We use the same methodology as the previous section, but selecting only the frequencies corresponding to periods of 6 to 32 quarters. The scaling factor is set to $c = \sqrt{0.13}$, which gives an acceptance rate of 25%.

The results are reported in the right panel of Table 7. A number of parameter estimates differ substantially from the full spectrum case. The most notable differences pertain to the parameters governing the exogenous shocks. Specifically, the AR coefficient of the total factor productivity process, ρ_a, drops from 0.98 to 0.78 while the standard deviation of its shock remains unchanged. The parameter governing the impact of productivity shocks on exogenous spending, ρ_{ga}, is almost

halved from 0.48 to 0.27. Additionally, the AR coefficient of the wage mark-up process ρ_w comes down from 0.94 to 0.63 and its MA coefficient μ_w drops from 0.84 to 0.36. The standard deviation of its shock decreases but the two posterior intervals overlap. On the other hand, the AR coefficients for risk premium (ρ_b) and monetary policy (ρ_r) shocks rise from 0.25 to 0.69, and from 0.15 to 0.36 respectively. The standard deviations of the respective shocks decrease from 0.24 and 0.24 to 0.10 and 0.13, respectively. The parameter differences outlined above are significant in the sense that their 90% probability intervals do not overlap. For the remaining three shock processes, exogenous spending, monetary policy and price mark-up, the magnitudes of the AR and MA coefficients either remain the same or show a small decrease, while the standard deviations of these shocks become smaller. Other notable differences in estimated parameters include the adjustment cost elasticity (φ), which goes down to 3.16 from 5.43, the degree of price indexation (ι_p), which increases from 0.31 to 0.59, and the coefficient on the lagged interest rate (ρ), which goes down from 0.82 to 0.74. These results imply that the model estimated using business cycle frequencies will potentially deliver different impulse responses from those obtained using the full spectrum. We explore this issue in the next section.

5 Impulse response analysis

Motivated by the differences found between parameter estimates obtained using the full spectrum and business cycle frequencies, we estimate the impulse response functions of the seven observables to the shocks for the two cases. Figures 2(a) through 2(g) report the posterior means, along with the 90% posterior intervals for horizons of up to 20 quarters. Each figure contains the response of a single observable to the seven shocks.

One notable difference between the responses of nearly all of the variables to a risk premium shock is that the impulse responses obtained using business cycle frequencies display a hump shaped dynamic, as opposed to an almost monotonic decay of those obtained using the full spectrum, as well as those in SW (2007). One exception is wage, where the impulse response in the full spectrum case is itself hump shaped, but still the pattern is much more pronounced in the business cycle frequency case.

The effects of both exogenous spending and investment shocks are significantly less pronounced when business cycle frequencies are used for estimation, with the exception of an investment shock to inflation and an exogenous spending shock to consumption and wage, for which the differences are not as clear cut.

The effect of a wage mark-up shock dies out faster for all variables if estimated using business cycle frequencies. Its effects are also significantly less pronounced after about 5 quarters for consumption and wage, after 10 quarters for output and labor hours, and for the whole 20 quarters for inflation and interest rate. It is interesting to note that in the business cycle case, the impulse response of investment to the wage mark-up shock is more pronounced initially for about five quarters, but then goes to zero faster after about 14 quarters.

The monetary policy shock also has a smaller impact and goes to zero faster in the business cycle case. Little difference can be observed when considering the responses to the price mark-up shock, as the two posterior intervals mostly overlap for the whole 20 quarters. However, responses become less pronounced and decay faster for consumption after roughly 10 quarters, and for output and labor hours after 15 quarters. The responses to the productivity shock are also very similar in the two cases, except for output, consumption and wage, for which the response is lower and decaying faster with business cycle frequencies.

It is important to ask whether the above difference is due to the impact of the prior, which has a greater effect in the business cycle frequency case as some information from the data is discarded. We address this as follows. First, we construct the log-likelihood using the business cycle frequencies, but evaluate it at the parameter values estimated from the full spectrum. Second, we construct the same likelihood function and evaluate it at the estimates from business cycle frequencies. The results are reported in Table 8. If the difference in the parameter estimates were entirely driven by the prior, then the likelihood in the second case would be smaller or of similar magnitude as the first case. The result suggests otherwise. Similarly, we construct the log-likelihood function using the full spectrum and evaluate it at both the business cycle and full spectrum based estimates. The difference is even more pronounced. Overall, the result suggests that business cycle based estimates achieve a better fit at such frequencies, but are at odds with other frequencies, in this case the very low frequencies as made clear below.

Since the above analysis omits frequencies from both sides of the business cycle frequency band, it leaves unclear which components are driving the difference. To investigate this, we first consider estimation omitting only frequencies below the business cycle band. Figures 3(a) to 3(g) contain the impulse responses for this case. The estimates from the full spectrum case are also included so that one can contrast 3(a)-3(g) with 2(a)-2(g). The figures show that the impulse responses computed omitting the low frequencies are overall similar to those using only business cycle frequencies. There are a few deviations from this pattern. The hump shaped responses of all seven variables to the risk

premium shock observed in business cycle results are no longer present. The same can be noted about the initial few quarters of responses of inflation to the productivity and the price mark-up shocks, as well as of wage to the price mark-up shock.

Next, we consider estimation omitting frequencies above the business cycle band. The results are reported in Figures 4(a) to 4(g), where again the full spectrum based estimates are included as the benchmark. Interestingly, the responses with high frequencies omitted are nearly identical to those estimated using the full spectrum. The few exceptions are that in the former case we observe the hump shaped response to the risk premium shock, a somewhat lower initial response to exogenous spending, investment, and productivity shocks of some variables, as well as lower initial responses of inflation and interest rate to the price mark-up and monetary policy shocks.

In summary, most of the differences observed between the impulse responses computed using the full spectrum estimates and those using business cycle frequencies can be attributed to the omission of the frequencies below the business cycle band.

6 Model diagnostics from a frequency domain perspective

King and Watson (1996) compared the spectra of three quantitative rational expectations models with that of the data. The models were calibrated and of small scale. Below, we carry out similar analysis for the medium scale DSGE model considered here. The goal of the analysis is two-fold. First, we examine whether the model captures the variability of and the comovements between relevant macroeconomic variables. Second, we compare the model spectrum estimated using all frequencies with that using only business cycle frequencies. The latter will highlight the potential value from using a subset of frequencies in estimation.

We obtain a nonparametric estimate of the spectral density by smoothing the periodograms using demeaned data. Suppose Y_t contains only one variable. Then, the estimator is given by

$$\widehat{f}(\omega_j) = \sum_{|k|\leq m} \mathcal{W}_T(k) I_T(\omega_{j+k}) \quad \text{for } j \geq 1$$

and

$$\widehat{f}(0) = \mathcal{W}_T(0) I_T(\omega_1) + 2\sum_{k=1}^{m} \mathcal{W}_T(k) I_T(\omega_{j+k}),$$

where m is a positive integer, $\mathcal{W}_T(k)$ is a weight function satisfying $\mathcal{W}_T(k) = \mathcal{W}_T(-k), \mathcal{W}_T(k) \geq 0$ $\forall$ k, $\sum_{|k|\leq m} \mathcal{W}_T(k) = 1$ and $I_T(\omega_j)$ is the periodogram. The estimator is consistent under mild conditions (see Brockwell and Davis (1991) for a rigorous treatment) and the asymptotic 95%

confidence intervals for the log of the spectral density are given by

$$\log(\widehat{f}(\omega_j)) \pm 1.96 \left(\sum_{|k| \leq m} \mathcal{W}_T(k)^2 \right)^{1/2}.$$

We apply the same type of estimator to obtain absolute coherency between pairs of variables. Let Y_t be a bivariate demeaned time series. The spectral density matrix is estimated in the same way as above but with $I_T(\omega_{j+k})$ being a 2×2 matrix. Let $\widehat{f}_{hk}(\omega_j)$ denote the (h,k)-th element of $\widehat{f}(\omega)$, then the absolute coherency estimate ($|\widehat{\mathfrak{K}}_{12}(\omega_j)|$) between Y_{1t} and Y_{2t} is

$$|\widehat{\mathfrak{K}}_{12}(\omega_j)| = [\widehat{c}_{12}^2(\omega_j) + \widehat{q}_{12}^2(\omega_j)]^{1/2} / [\widehat{f}_{11}(\omega_j)\widehat{f}_{22}(\omega_j)]^{1/2},$$

where

$$\begin{aligned} \widehat{c}_{12}(\omega_j) &= [\widehat{f}_{12}(\omega_j) + \widehat{f}_{21}(\omega_j)]/2, \\ \widehat{q}_{12}(\omega_j) &= i[\widehat{f}_{12}(\omega_j) - \widehat{f}_{21}(\omega_j)]/2. \end{aligned}$$

The 95% confidence intervals can be computed as

$$|\widehat{\mathfrak{K}}_{12}(\omega_j)| \ \pm 1.96(1 - |\widehat{\mathfrak{K}}_{12}(\omega_j)|^2) \left(\sum_{|k| \leq m} \mathcal{W}_T(k)^2 \right)^{1/2} / \sqrt{2}.$$

In applications, the choice of $\mathcal{W}_T(k)$ depends on the characteristics of the data series at hand. It is possible and sometimes advantageous to use different weighting functions for estimation of different elements of the spectral density matrix due to potentially different features of the time series (see Ch. 9 in Priestley (1981) for a discussion). In our case, we apply the same weight function in all estimations, with $m = 4$ and the weights given by $\left\{ \frac{1}{21}, \frac{2}{21}, \frac{3}{21}, \frac{3}{21}, \frac{3}{21}, \frac{3}{21}, \frac{3}{21}, \frac{2}{21}, \frac{1}{21} \right\}$, which is obtained by the successive application of two Daniell filters with weights given by $\left\{ \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right\}$ and $\left\{ \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7} \right\}$. This choice of $\mathcal{W}_T(k)$ produces spectra estimates that are not as rough as the raw periodogram, and in the meantime do not appear oversmoothed.

Figure 5 plots the log spectra of the seven variables. Three results are reported in each subfigure. First, we report the nonparametric estimates of the spectrum of the demeaned data series along with the pointwise 95% confidence intervals. They are used as a benchmark to assess the model's ability in capturing these key features. The solid curve is the spectrum implied by the model with parameters estimated using the full spectrum. The dashed line is the same object but with business cycle based estimates. Two patterns emerge. First, the solid curve captures the overall shape of the data spectrum, although there are noticeable departures which often occur inside of

the business cycle frequencies. It should be noted that for the growth series (sub-figures i-iii,vi), the model implies that their spectral density at frequency zero is zero (as the figure reports log spectra, the frequency zero is omitted from the figures). This is inconsistent with the data spectra, which are positive at the origin. When frequencies very near zero are included in the estimation, the model will try to reduce such a departure by having very persistent estimates. This potentially affects the other frequencies, which partly explains why the full spectrum based estimates do not capture the slope of the spectrum very well inside of the business cycle frequencies. When using only business cycle frequencies for estimation, such a tension is absent and the estimates do a better job at matching variations at these frequencies. The lines never fall substantially outside of the confidence bands based on the nonparametric estimates. However, the departures from the data spectrum can be substantial outside of the business cycle frequencies. In practice, this offers the researcher a choice. If one firmly believes that the DSGE model is well specified at all frequencies, then, they should all enter the estimation and the estimates will be more efficient. If one suspects that the modeling of the trend, or, more generally, of the very low frequency behavior in the model is inconsistent with the data (for example, the data has a broken trend while the model has a linear trend), then the subset based approach may be a more robust choice.

Figures 6(a) to 6(c) report the absolute coherency between the seven variables. Notice that their values can be interpreted as a measure of strength of correlation at a particular frequency. Both the business cycle and the full spectrum based estimates achieve something at capturing their overall magnitudes, with the exception of the comovements between interest rate and other four variables (consumption growth, investment growth, output growth, and labor hours). In the latter case, the two estimates are close and are consistently below the nonparametric estimates. This unanimous finding suggests a dimension along which the model can be further improved. For the other cases, the business cycle based estimates typically do a better job at the intended frequencies. They largely stay within the confidence intervals, and are better at capturing the peaks of the coherency, while the full spectrum based estimates miss them in the majority of cases.

In summary, the SW(2007) model does a reasonable job at matching the spectra of individual time series and the absolute coherency implied by the data. The business cycle based estimates offer the flexibility to focus on a particular frequency band and to achieve a better fit at such frequencies. In practice, both analyses can be carried out, allowing us to assess to what extent the results are driven by some particular frequencies.

7 Conclusion and discussion

The paper has considered identification, estimation and inference in medium scale DSGE models using SW (2007) as an illustrative example. A key element in the analysis is that we can focus on part of the spectrum.

For identification, we derived the non-identification curve to reveal which and how many parameters need to be fixed to achieve local identification. For estimation and inference, we compared estimates obtained using the full spectrum with those using only business cycle frequencies and reported notably different parameter values and impulse response functions. Further analysis shows that the differences are mainly due to the frequencies below the business cycle frequency band. We have also considered model diagnostics by contrasting the model based and the nonparametrically estimated spectra as well as examining the absolute coherency. The result suggests that SW (2007) does a reasonable job at matching these two features observed in the data, with the exception of the comovements between interest rate and other four variables (consumption growth, investment growth, output growth, and labor hours). The business cycle based estimate, due to its ability to focus on a particular frequency band, achieves a better fit at such frequencies.

From a methodological perspective, the results contribute to the relatively sparse literature that exploits the advantage of model estimation and diagnostics using a subset of frequencies. Engle (1974) is a seminal contribution. It proposed band spectrum regression as a way to allow for frequency specific misspecification, seasonality and measurement errors, and to obtain better understanding of some common time domain procedures, such as applying a moving average filter. Sims (1993) and Hansen and Sargent (1993) considered the effect of removing or downweighting seasonal frequencies on estimating rational expectations models. Diebold, Ohanian and Berkowitz (1998) discussed a general framework for loss function based estimation and model evaluation. In a different context, McCloskey (2010) considered parameter estimation in ARMA, GARCH and stochastic volatility models robust to low frequency contamination caused by a changing mean or misspecified trend. Qu and Tkachenko (2012) provided a comprehensive treatment of the theoretical and computational aspects of the frequency domain quasi-likelihood applied to DSGE models. By working through a concrete example, this paper demonstrates that such an approach is applicable to medium scale DSGE models and that it offers substantial depth and flexibility when compared with time domain methods. We intend to apply the methodology to a relatively broad class of DSGE models and hope to report results in the near future.

References

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

Brockwell, P., and R. Davis (1991): *Time Series: Theory and Methods*. New York: Springer-Verlag, 2 edn.

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

Chernozhukov, V., and H. Hong (2003): "An MCMC Approach to Classical Estimation," *Journal of Econometrics*, 115, 293–346.

Diebold, F. X., L. E. Ohanian, and J. Berkowitz (1998): "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data," *The Review of Economic Studies*, 65, 433–451.

Engle, R. F. (1974): "Band Spectrum Regression," *International Economic Review*, 15, 1–11.

Hansen, L. P., and T. J. Sargent (1993): "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics*, 55, 21–55.

Hastings, W. K. (1970): "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57, 97–109.

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

King, R. G., and M. W. Watson (1996): "Money, Prices, Interest Rates and the Business Cycle," *The Review of Economics and Statistics*, 78, 35–53.

Komunjer, I., and S. Ng (2011): "Dynamic Identification of DSGE Models," *Econometrica*, 79, 1995–2032.

Leeper, E. M., and C. A. Sims (1994): "Toward a Modern Macroeconomic Model Usable for Policy Analysis," in *NBER Macroeconomics Annual*, ed. by F. Stanley, and J. J. Rotemberg, pp. 81–118. MIT Press, Cambridge, MA.

McCloskey, A. (2010): "Parameter Estimation Robust to Low-Frequency Contamination with Applications to ARMA, GARCH and Stochastic Volatility Models," Working paper, Boston University.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953): "Equation of State Calculations by Fast Computing Machines," *Journal of Chemical Physics*, 21, 1087–1092.

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

Priestley, M. B. (1981): *Spectral Analysis and Time Series*, vol. 1. New York: Academic Press.

Qu, Z., and D. Tkachenko (2012): "Identification and Frequency Domain Quasi-Maximum Likelihood Estimation of Linearized Dynamic Stochastic General Equilibrium Models," *Quantitative Economics*, 3, 95–132.

Roberts, G., A. Gelman, and W. Gilks (1997): "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms," *Annals of Applied Probability*, 7, 110–120.

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

——— (2011): "Estimation and Evaluation of DSGE Models: Progress and Challenges," NBER Working Paper 16781.

Sims, C. A. (1993): "Rational Expectations Modeling with Seasonally Adjusted Data," *Journal of Econometrics*, 55, 9–19.

——— (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20, 1–20.

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

Uhlig, H. (1999): "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott. Oxford University Press.

Watson, M. W. (1993): "Measures of Fit for Calibrated Models," *Journal of Political Economy*, 101, 1011–1041.

Appendix: Equations and variables in the flexible price-wage economy

The equations are similar to those in the sticky price-wage economy, but with the variables μ_t^p and μ_t^w set to zero. Stars are used to denote variables from this economy. Because the shock processes are identical to those in the sticky price-wage economy, we do not repeat them here.

1. The resource constraint:
$$y_t^* = c_y c_t^* + i_y i_t^* + z_y z_t^* + \varepsilon_t^g.$$

2. The dynamics of consumption follow from the consumption Euler equation
$$c_t^* = c_1 c_{t-1}^* + (1-c_1) E_t c_{t+1}^* + c_2 (l_t^* - E_t l_{t+1}^*) - c_3 (r_t^* - 0) - \varepsilon_t^b.$$
Note that the expected inflation is zero because the price adjusts instantaneously.

3. The dynamics of investment come from the investment Euler equation
$$i_t^* = i_1 i_{t-1}^* + (1-i_1)\, E_t i_{+1}^* + i_2 q_t^* + \varepsilon_t^i.$$

4. The corresponding arbitrage equation for the value of capital is given by
$$q_t^* = q_1 E_t q_{t+1}^* + (1-q_1)\, E_t r_{t+1}^{*k} - (r_t^* - 0) - \frac{1}{c_3}\varepsilon_t^b.$$
The expected inflation is zero for the same reason as above.

5. The aggregate production function is
$$y_t^* = \phi_p \left(\alpha k_t^{*s} + (1-\alpha)\, l_t^* + \varepsilon_t^a\right).$$

6. Current capital service use is a function of capital installed in the previous period and the degree of capital utilization
$$k_t^{*s} = k_{t-1}^* + z_t^*.$$

7. The degree of capital utilization is a positive fraction of the rental rate of capital
$$z_t^* = z_1 r_t^{*k}.$$

8. The accumulation of installed capital is
$$k_t^* = k_1 k_{t-1}^* + (1-k_1)\, i_t^* + k_2 \varepsilon_t^i.$$

9. Because $\mu_t^p = 0$ and the relationship with rigidity is:
$$\mu_t^p = \alpha\,(k_t^s - l_t) + \varepsilon_t^a - w_t,$$
we have
$$0 = \alpha\,(k_t^s - l_t) + \varepsilon_t^a - w_t$$
or, equivalently,
$$\alpha r_t^{*k} + (1-\alpha)\, w_t^* = \varepsilon_t^a.$$
There is no New Keynesian Phillips curve as price adjusts instantaneously.

10. The rental rate of capital is

$$r_t^{*k} = -\left(k_t^{*s} - l_t^*\right) + w_t^*.$$

11. The wage mark-up is now $\mu_t^w = 0$. Therefore,

$$0 = w_t^* - \left(\sigma_l l_t^* + \frac{1}{1-\lambda}\left(c_t^* - \lambda c_{t-1}^*\right)\right)$$

or

$$w_t^* = \left(\sigma_l l_t^* + \frac{1}{1-\lambda}\left(c_t^* - \lambda c_{t-1}^*\right)\right).$$

Table 1: Parameter values and the corresponding two smallest eigenvalues along the non-identification curve

	φ	λ	γ	β	δ	λ_1	λ_2
θ_0	5.740000	0.710000	1.004300	0.998400	0.025000	1.80E-10	0.392865
				Panel (a). Direction 1			
θ_1	12.417476	0.482721	0.682812	0.862248	0.337109	1.96E-14	0.808082
θ_2	19.113813	0.389080	0.550356	0.794406	0.465700	4.57E-14	1.210705
θ_3	25.812574	0.334809	0.473589	0.750327	0.540228	3.01E-14	1.599268
θ_4	32.512006	0.298325	0.42198	0.718141	0.590328	5.53E-15	1.975594
θ_5	39.211698	0.271647	0.384246	0.693026	0.626964	3.26E-15	2.341212
θ_6	45.911511	0.25105	0.35511	0.672563	0.655256	8.58E-15	2.697239
θ_7	52.611389	0.234516	0.331724	0.655380	0.677954	1.04E-14	3.044732
θ_8	59.311305	0.220873	0.312427	0.640622	0.696688	4.10E-15	3.384357
θ_9	66.011244	0.209364	0.296147	0.627727	0.712493	5.40E-15	3.716722
θ_{10}	72.711198	0.199485	0.282174	0.616303	0.726059	9.96E-16	4.042423
				Panel (b). Direction 2			
θ_{-1}	5.735346	0.710288	1.004707	0.998556	0.024605	5.27E-12	0.392485
θ_{-2}	5.730692	0.710576	1.005115	0.998711	0.024209	3.00E-12	0.392186
θ_{-3}	5.726038	0.710865	1.005523	0.998865	0.023812	2.95E-11	0.391895
θ_{-4}	5.721384	0.711154	1.005933	0.999019	0.023415	3.93E-11	0.391616
θ_{-5}	5.716730	0.711444	1.006342	0.999173	0.023018	9.91E-11	0.391323
θ_{-6}	5.712077	0.711732	1.006752	0.999328	0.022620	1.12E-10	0.391078
θ_{-7}	5.707423	0.712023	1.007162	0.999483	0.022221	8.78E-11	0.390749
θ_{-8}	5.702770	0.712314	1.007573	0.999638	0.021823	8.39E-11	0.390467
θ_{-9}	5.698117	0.712605	1.007984	0.999793	0.021423	1.97E-10	0.390278
θ_{-10}	5.693464	0.712896	1.008396	0.999948	0.021024	1.13E-10	0.389814

Note. θ_j represent equally spaced points taken from the non-identification curve extended from θ_0 for 670,000 steps in Direction 1, and for 472 steps in Direction 2. λ_1 and λ_2 represent the smallest and the second smallest eigenvalues of $G(\theta_i)^s$. The step size for computing the curve is 10^{-4}. Along Direction 1, the curve is truncated at the point where β is closest to 1, as it is the discount factor. Results are rounded to the nearest sixth digit to the right of decimal.

Table 2: Deviations of spectra across frequencies (Direction 1)

	10 largest deviations across frequencies and elements ordered in descending order									
	1	2	3	4	5	6	7	8	9	10
	Maximum absolute deviations across frequencies									
θ_1	8.99E-05	2.98E-05	1.24E-05	1.24E-05	1.09E-05	1.09E-05	9.24E-06	9.24E-06	5.97E-06	5.97E-06
θ_2	1.17E-04	3.88E-05	1.61E-05	1.61E-05	1.42E-05	1.42E-05	1.20E-05	1.20E-05	7.77E-06	7.77E-06
θ_3	1.31E-04	4.31E-05	1.79E-05	1.79E-05	1.59E-05	1.59E-05	1.33E-05	1.33E-05	8.65E-06	8.65E-06
θ_4	1.38E-04	4.57E-05	1.90E-05	1.90E-05	1.68E-05	1.68E-05	1.41E-05	1.41E-05	9.16E-06	9.16E-06
θ_5	1.43E-04	4.74E-05	1.97E-05	1.97E-05	1.74E-05	1.74E-05	1.46E-05	1.46E-05	9.50E-06	9.50E-06
θ_6	1.47E-04	4.85E-05	2.02E-05	2.02E-05	1.78E-05	1.78E-05	1.50E-05	1.50E-05	9.72E-06	9.72E-06
θ_7	1.49E-04	4.94E-05	2.05E-05	2.05E-05	1.81E-05	1.81E-05	1.53E-05	1.53E-05	9.89E-06	9.89E-06
θ_8	1.51E-04	5.01E-05	2.08E-05	2.08E-05	1.83E-05	1.83E-05	1.55E-05	1.55E-05	1.00E-05	1.00E-05
θ_9	1.52E-04	5.06E-05	2.10E-05	2.10E-05	1.84E-05	1.84E-05	1.56E-05	1.56E-05	1.01E-05	1.01E-05
θ_{10}	1.53E-04	5.10E-05	2.12E-05	2.12E-05	1.86E-05	1.86E-05	1.58E-05	1.58E-05	1.02E-05	1.02E-05
	Maximum absolute deviations across frequencies in relative form									
θ_1	7.81E-06	5.33E-06	4.60E-06	4.34E-06	4.19E-06	3.73E-06	3.73E-06	3.34E-06	3.34E-06	2.91E-06
θ_2	1.02E-05	6.93E-06	5.98E-06	5.66E-06	5.47E-06	4.86E-06	4.86E-06	4.39E-06	4.34E-06	3.80E-06
θ_3	1.13E-05	7.73E-06	6.70E-06	6.32E-06	6.13E-06	5.41E-06	5.41E-06	4.89E-06	4.84E-06	4.23E-06
θ_4	1.20E-05	8.18E-06	7.07E-06	6.69E-06	6.49E-06	5.73E-06	5.73E-06	5.18E-06	5.13E-06	4.48E-06
θ_5	1.24E-05	8.48E-06	7.31E-06	6.93E-06	6.71E-06	5.93E-06	5.93E-06	5.37E-06	5.31E-06	4.64E-06
θ_6	1.27E-05	8.68E-06	7.50E-06	7.09E-06	6.86E-06	6.07E-06	6.07E-06	5.50E-06	5.43E-06	4.75E-06
θ_7	1.29E-05	8.82E-06	7.63E-06	7.21E-06	6.96E-06	6.18E-06	6.18E-06	5.59E-06	5.52E-06	4.84E-06
θ_8	1.31E-05	8.91E-06	7.71E-06	7.28E-06	7.07E-06	6.25E-06	6.25E-06	5.61E-06	5.60E-06	4.90E-06
θ_9	1.33E-05	8.98E-06	7.79E-06	7.34E-06	7.12E-06	6.31E-06	6.31E-06	5.66E-06	5.65E-06	4.95E-06
θ_{10}	1.34E-05	9.05E-06	7.85E-06	7.40E-06	7.18E-06	6.36E-06	6.36E-06	5.71E-06	5.70E-06	4.99E-06
	Maximum relative deviations across frequencies									
θ_1	5.94E-05	5.94E-05	2.67E-05	2.67E-05	1.52E-05	1.52E-05	1.37E-05	1.37E-05	8.46E-06	7.25E-06
θ_2	7.75E-05	7.75E-05	3.49E-05	3.49E-05	1.99E-05	1.99E-05	1.79E-05	1.79E-05	1.10E-05	9.43E-06
θ_3	8.65E-05	8.65E-05	3.91E-05	3.91E-05	2.23E-05	2.23E-05	2.00E-05	2.00E-05	1.22E-05	1.05E-05
θ_4	9.16E-05	9.16E-05	4.14E-05	4.14E-05	2.36E-05	2.36E-05	2.12E-05	2.12E-05	1.30E-05	1.11E-05
θ_5	9.48E-05	9.48E-05	4.28E-05	4.28E-05	2.44E-05	2.44E-05	2.20E-05	2.20E-05	1.34E-05	1.15E-05
θ_6	9.71E-05	9.71E-05	4.38E-05	4.38E-05	2.49E-05	2.49E-05	2.24E-05	2.24E-05	1.38E-05	1.18E-05
θ_7	9.88E-05	9.88E-05	4.45E-05	4.45E-05	2.53E-05	2.53E-05	2.28E-05	2.28E-05	1.40E-05	1.20E-05
θ_8	9.99E-05	9.99E-05	4.49E-05	4.49E-05	2.55E-05	2.55E-05	2.30E-05	2.30E-05	1.42E-05	1.21E-05
θ_9	1.01E-04	1.01E-04	4.54E-05	4.54E-05	2.56E-05	2.56E-05	2.32E-05	2.32E-05	1.44E-05	1.23E-05
θ_{10}	1.02E-04	1.02E-04	4.58E-05	4.58E-05	2.58E-05	2.58E-05	2.34E-05	2.34E-05	1.45E-05	1.24E-05

Note. θ_1 to θ_{10} are as defined in Table 1. We report 10 largest deviations across 49 elements of each $G(\theta_i)^s$ computed at 5,000 frequencies to conserve space.

Table 3: Deviations of spectra across frequencies (Direction 2)

	10 largest deviations across frequencies and elements ordered in descending order									
	1	2	3	4	5	6	7	8	9	10
	Maximum absolute deviations across frequencies									
θ_{-1}	1.59E-07	3.14E-08	3.14E-08	3.09E-08	1.77E-08	1.77E-08	1.65E-08	1.65E-08	1.09E-08	1.09E-08
θ_{-2}	2.38E-07	6.50E-08	4.93E-08	4.93E-08	2.33E-08	2.33E-08	2.27E-08	2.27E-08	2.22E-08	2.22E-08
θ_{-3}	3.54E-07	1.14E-07	5.52E-08	5.52E-08	3.80E-08	3.80E-08	2.59E-08	2.59E-08	2.19E-08	2.19E-08
θ_{-4}	5.88E-07	1.68E-07	8.09E-08	8.09E-08	7.20E-08	7.20E-08	5.26E-08	5.26E-08	3.75E-08	3.75E-08
θ_{-5}	8.55E-07	2.42E-07	1.12E-07	1.12E-07	1.07E-07	1.07E-07	8.22E-08	8.22E-08	5.54E-08	5.54E-08
θ_{-6}	1.08E-06	3.11E-07	1.34E-07	1.34E-07	1.24E-07	1.24E-07	8.90E-08	8.90E-08	6.83E-08	6.83E-08
θ_{-7}	1.32E-06	3.76E-07	1.82E-07	1.82E-07	1.54E-07	1.54E-07	1.39E-07	1.39E-07	8.60E-08	8.60E-08
θ_{-8}	1.40E-06	4.11E-07	1.83E-07	1.83E-07	1.62E-07	1.62E-07	1.30E-07	1.30E-07	9.02E-08	9.02E-08
θ_{-9}	1.44E-06	4.42E-07	1.80E-07	1.80E-07	1.62E-07	1.62E-07	1.18E-07	1.18E-07	9.18E-08	9.18E-08
θ_{-10}	1.47E-06	4.57E-07	1.80E-07	1.80E-07	1.71E-07	1.71E-07	1.17E-07	1.17E-07	9.35E-08	9.35E-08
	Maximum absolute deviations across frequencies in relative form									
θ_{-1}	2.24E-08	1.54E-08	1.36E-08	1.16E-08	1.07E-08	9.44E-09	7.97E-09	7.78E-09	6.30E-09	6.30E-09
θ_{-2}	4.52E-08	3.14E-08	2.63E-08	1.82E-08	1.65E-08	1.56E-08	1.51E-08	1.51E-08	1.50E-08	1.30E-08
θ_{-3}	4.03E-08	2.96E-08	2.82E-08	2.75E-08	1.81E-08	1.48E-08	1.45E-08	1.45E-08	1.44E-08	1.34E-08
θ_{-4}	4.37E-08	4.35E-08	3.55E-08	3.55E-08	3.00E-08	2.52E-08	2.30E-08	2.30E-08	2.13E-08	2.12E-08
θ_{-5}	1.07E-07	6.35E-08	4.97E-08	4.44E-08	3.98E-08	3.56E-08	3.56E-08	3.09E-08	3.04E-08	2.86E-08
θ_{-6}	1.50E-07	8.22E-08	5.96E-08	5.90E-08	5.21E-08	4.62E-08	4.62E-08	4.11E-08	3.85E-08	3.68E-08
θ_{-7}	1.69E-07	9.95E-08	7.35E-08	7.21E-08	5.73E-08	5.73E-08	5.08E-08	4.82E-08	4.70E-08	4.55E-08
θ_{-8}	1.81E-07	1.08E-07	7.71E-08	7.59E-08	6.21E-08	6.02E-08	6.02E-08	5.22E-08	5.05E-08	4.74E-08
θ_{-9}	1.87E-07	1.17E-07	7.91E-08	7.73E-08	7.12E-08	6.13E-08	6.13E-08	5.62E-08	5.16E-08	4.85E-08
θ_{-10}	1.91E-07	1.20E-07	8.17E-08	7.94E-08	7.69E-08	6.10E-08	6.10E-08	5.76E-08	5.35E-08	5.29E-08
	Maximum relative deviations across frequencies									
θ_{-1}	8.38E-08	8.38E-08	6.39E-08	6.39E-08	5.12E-08	5.12E-08	3.22E-08	3.22E-08	2.21E-08	2.21E-08
θ_{-2}	2.51E-07	2.51E-07	1.38E-07	1.38E-07	1.23E-07	1.23E-07	5.72E-08	5.72E-08	4.92E-08	4.92E-08
θ_{-3}	3.32E-07	3.32E-07	1.68E-07	1.68E-07	1.12E-07	1.12E-07	7.00E-08	7.00E-08	3.72E-08	3.72E-08
θ_{-4}	3.76E-07	3.76E-07	1.89E-07	1.89E-07	1.39E-07	1.39E-07	1.02E-07	1.02E-07	5.18E-08	5.18E-08
θ_{-5}	4.58E-07	4.58E-07	2.23E-07	2.23E-07	1.64E-07	1.64E-07	1.42E-07	1.42E-07	7.58E-08	7.58E-08
θ_{-6}	6.72E-07	6.72E-07	3.34E-07	3.34E-07	2.28E-07	2.28E-07	1.93E-07	1.93E-07	1.21E-07	1.21E-07
θ_{-7}	6.52E-07	6.52E-07	3.07E-07	3.07E-07	2.63E-07	2.63E-07	2.14E-07	2.14E-07	1.82E-07	1.82E-07
θ_{-8}	8.18E-07	8.18E-07	3.95E-07	3.95E-07	2.78E-07	2.78E-07	2.38E-07	2.38E-07	1.63E-07	1.63E-07
θ_{-9}	9.84E-07	9.84E-07	4.79E-07	4.79E-07	2.88E-07	2.88E-07	2.55E-07	2.55E-07	1.41E-07	1.41E-07
θ_{-10}	1.06E-06	1.06E-06	5.19E-07	5.19E-07	2.97E-07	2.97E-07	2.62E-07	2.62E-07	1.30E-07	1.30E-07

Note. θ_{-1} to θ_{-10} are as defined in Table 1. We report 10 largest deviations across 49 elements of each $G(\theta_i)^s$ computed at 5,000 frequencies to conserve space.

Table 4: Rank sensitivity analysis

	Differentiation step size $\times \theta_0$							
	1E-02	1E-03	1E-04	1E-05	1E-06	1E-07	1E-08	1E-09
TOL	Rank of $G(\theta_0)$							
1E-03	37	36	36	36	36	36	36	36
1E-04	37	37	37	36	36	36	36	36
1E-05	37	37	37	36	36	36	36	36
1E-06	37	37	37	36	36	36	36	36
1E-07	38	37	37	37	36	36	36	37
1E-08	39	37	37	37	36	36	37	37
1E-09	39	38	38	37	37	36	37	37
1E-10	39	39	39	37	37	37	37	39
Default	39	38	37	37	37	36	37	37

Note. TOL refers to the tolerance level used to determine the rank. Default refers to the MATLAB default tolerance level.

Table 5: Prior distributions of the parameters

	Parameter interpretation	Distribution	Mean	St. Dev.
ρ_{ga}	Cross-corr.: tech. and exog. spending shocks	Normal	0.50	0.25
μ_w	Wage mark-up shock MA	Beta	0.50	0.20
μ_p	Price mark-up shock MA	Beta	0.50	0.20
α	Share of capital in production	Normal	0.30	0.05
ψ	Elasticity of capital utilization adjustment cost	Beta	0.50	0.15
φ	Investment adjustment cost	Normal	4.00	1.50
σ_c	Elasticity of inertemporal substitution	Normal	1.50	0.38
λ	Habit persistence	Beta	0.70	0.10
ϕ_p	Fixed costs in production	Normal	1.25	0.13
ι_w	Wage indexation	Beta	0.50	0.15
ξ_w	Wage stickiness	Beta	0.50	0.10
ι_p	Price indexation	Beta	0.50	0.15
ξ_p	Price stickiness	Beta	0.50	0.10
σ_l	Labor supply elasticity	Normal	2.00	0.75
r_π	Taylor rule: inflation weight	Normal	1.50	0.25
$r_{\Delta y}$	Taylor rule: feedback from output gap change	Normal	0.13	0.05
r_y	Taylor rule: output gap weight	Normal	0.13	0.05
ρ	Taylor rule: interest rate smoothing	Beta	0.75	0.10
ρ_a	Productivity shock AR	Beta	0.50	0.20
ρ_b	Risk premium shock AR	Beta	0.50	0.20
ρ_g	Exogenous spending shock AR	Beta	0.50	0.20
ρ_i	Interest rate shock AR	Beta	0.50	0.20
ρ_r	Monetary policy shock AR	Beta	0.50	0.20
ρ_p	Price mark-up shock AR	Beta	0.50	0.20
ρ_w	Wage mark-up shock AR	Beta	0.50	0.20
σ_a	Productivity shock std. dev.	Invgamma	0.10	2.00
σ_b	Risk premium shock std. dev.	Invgamma	0.10	2.00
σ_g	Exogenous spending shock std. dev.	Invgamma	0.10	2.00
σ_i	Interest rate shock std. dev.	Invgamma	0.10	2.00
σ_r	Monetary policy shock std. dev.	Invgamma	0.10	2.00
σ_p	Price mark-up shock std. dev.	Invgamma	0.10	2.00
σ_w	Wage mark-up shock std. dev.	Invgamma	0.10	2.00
$\overline{\gamma}$	Trend growth rate: real GDP, Infl., Wages	Normal	0.40	0.10
$100(\beta^{-1}-1)$	Discount rate	Gamma	0.25	0.10
$\overline{\pi}$	Steady state inflation rate	Gamma	0.62	0.10
$\bar{l}$	Steady state hours worked	Normal	0.00	2.00
δ	Capital depreciation rate	Fixed	0.025	
ϕ_w	Steady state labor market mark-up	Fixed	1.50	
g_y	Steady state exog. spending-output ratio	Fixed	0.18	
ϵ_p	Curvature of Kimball goods market aggregator	Fixed	10.00	
ϵ_w	Curvature of Kimball labor market aggregator	Fixed	10.00	

Note. Prior distributions are taken from SW(2007) Dynare code.

Table 6: Posterior distributions of the parameters

	Full Spectrum and mean				SW(2007) Tables 1 A,B			
	Mode	Mean	5%	95%	Mode	Mean	5%	95%
ρ_{ga}	0.49	0.49	0.36	0.62	0.52	0.52	0.37	0.66
μ_w	0.90	0.86	0.75	0.93	0.88	0.84	0.75	0.93
μ_p	0.67	0.63	0.44	0.79	0.74	0.69	0.54	0.85
α	0.20	0.20	0.17	0.23	0.19	0.19	0.16	0.21
ψ	0.64	0.62	0.44	0.79	0.54	0.54	0.36	0.72
φ	5.14	5.21	3.85	6.68	5.48	5.74	3.97	7.42
σ_c	1.67	1.61	1.33	1.93	1.39	1.38	1.16	1.59
λ	0.67	0.68	0.60	0.75	0.71	0.71	0.64	0.78
ϕ_p	1.55	1.55	1.43	1.68	1.61	1.60	1.48	1.73
ι_w	0.54	0.54	0.33	0.74	0.59	0.58	0.38	0.78
ξ_w	0.77	0.74	0.63	0.83	0.73	0.70	0.60	0.81
ι_p	0.26	0.29	0.14	0.47	0.22	0.24	0.10	0.38
ξ_p	0.65	0.65	0.57	0.72	0.65	0.66	0.56	0.74
σ_l	1.80	1.66	0.74	2.72	1.92	1.83	0.91	2.78
r_π	2.00	2.02	1.75	2.31	2.03	2.04	1.74	2.33
$r_{\Delta y}$	0.22	0.23	0.18	0.27	0.22	0.22	0.18	0.27
r_y	0.11	0.11	0.07	0.16	0.08	0.08	0.05	0.12
ρ	0.83	0.82	0.78	0.86	0.81	0.81	0.77	0.85
ρ_a	0.98	0.98	0.97	0.99	0.95	0.95	0.94	0.97
ρ_b	0.22	0.25	0.11	0.41	0.18	0.22	0.07	0.36
ρ_g	0.92	0.91	0.86	0.95	0.97	0.97	0.96	0.99
ρ_i	0.72	0.73	0.63	0.82	0.71	0.71	0.61	0.80
ρ_r	0.12	0.15	0.06	0.26	0.12	0.15	0.04	0.24
ρ_p	0.84	0.82	0.72	0.91	0.90	0.89	0.80	0.96
ρ_w	0.96	0.95	0.91	0.98	0.97	0.96	0.94	0.99
σ_a	0.48	0.49	0.44	0.54	0.45	0.45	0.41	0.50
σ_b	0.24	0.24	0.20	0.28	0.24	0.23	0.19	0.27
σ_g	0.50	0.50	0.46	0.55	0.52	0.53	0.48	0.58
σ_i	0.46	0.46	0.39	0.54	0.45	0.45	0.37	0.53
σ_r	0.23	0.24	0.22	0.27	0.24	0.24	0.22	0.27
σ_p	0.15	0.15	0.12	0.17	0.14	0.14	0.11	0.16
σ_w	0.24	0.24	0.20	0.27	0.24	0.24	0.20	0.28
$\overline{\gamma}$	0.31	0.31	0.20	0.42	0.43	0.43	0.40	0.45
$100(\beta^{-1}-1)$	0.17	0.21	0.10	0.34	0.16	0.16	0.07	0.26
$\overline{\pi}$	0.68	0.69	0.54	0.87	0.81	0.78	0.61	0.96
$\bar{l}$	0.59	0.79	0.04	1.60	-0.1	0.53	-1.3	2.32

Note: 5% and 95% columns refer to the 5th and 95th percentiles of the distribution of RWM draws.

Table 7: Posterior distribution of the dynamic parameters

	Full Spectrum				Business cycle			
	Mode	Mean	5%	95%	Mode	Mean	5%	95%
ρ_{ga}	0.49	0.48	0.35	0.62	0.26	0.27	0.09	0.45
μ_w	0.90	0.84	0.74	0.92	0.29	0.36	0.12	0.63
μ_p	0.67	0.68	0.53	0.80	0.57	0.48	0.20	0.73
α	0.21	0.20	0.17	0.23	0.18	0.19	0.15	0.22
ψ	0.63	0.62	0.43	0.79	0.50	0.52	0.30	0.74
φ	5.13	5.43	3.83	7.21	2.57	3.16	2.14	4.71
σ_c	1.76	1.71	1.40	2.05	1.29	1.41	1.11	1.80
λ	0.65	0.66	0.58	0.74	0.60	0.59	0.47	0.70
ϕ_p	1.56	1.56	1.43	1.69	1.43	1.44	1.31	1.59
ι_w	0.54	0.56	0.36	0.76	0.51	0.53	0.29	0.75
ξ_w	0.77	0.75	0.64	0.84	0.75	0.74	0.64	0.82
ι_p	0.26	0.31	0.17	0.46	0.60	0.59	0.33	0.81
ξ_p	0.65	0.64	0.57	0.71	0.67	0.66	0.58	0.75
σ_l	1.70	1.59	0.70	2.57	2.23	2.07	1.02	3.17
r_π	2.02	2.02	1.75	2.31	1.93	1.93	1.64	2.24
$r_{\Delta y}$	0.22	0.22	0.18	0.27	0.20	0.20	0.16	0.25
r_y	0.11	0.11	0.07	0.15	0.14	0.14	0.08	0.20
ρ	0.83	0.82	0.78	0.86	0.74	0.74	0.68	0.80
ρ_a	0.98	0.98	0.97	0.99	0.79	0.78	0.61	0.92
ρ_b	0.21	0.25	0.11	0.40	0.76	0.69	0.51	0.84
ρ_g	0.92	0.91	0.86	0.95	0.89	0.87	0.77	0.95
ρ_i	0.73	0.74	0.65	0.83	0.70	0.66	0.48	0.81
ρ_r	0.12	0.15	0.05	0.26	0.39	0.36	0.14	0.58
ρ_p	0.85	0.85	0.76	0.92	0.78	0.73	0.50	0.90
ρ_w	0.96	0.94	0.89	0.98	0.63	0.63	0.41	0.80
σ_a	0.48	0.48	0.44	0.53	0.47	0.49	0.40	0.59
σ_b	0.24	0.24	0.20	0.28	0.08	0.10	0.06	0.14
σ_g	0.50	0.50	0.46	0.55	0.35	0.37	0.31	0.45
σ_i	0.46	0.45	0.39	0.53	0.32	0.38	0.25	0.58
σ_r	0.24	0.24	0.22	0.27	0.12	0.13	0.10	0.18
σ_p	0.15	0.15	0.12	0.18	0.08	0.08	0.05	0.12
σ_w	0.24	0.24	0.20	0.27	0.15	0.19	0.11	0.31
$\overline{\gamma}$	0.41	0.42	0.27	0.59	0.40	0.40	0.24	0.57
$100(\beta^{-1}-1)$	0.22	0.26	0.12	0.45	0.23	0.27	0.12	0.46

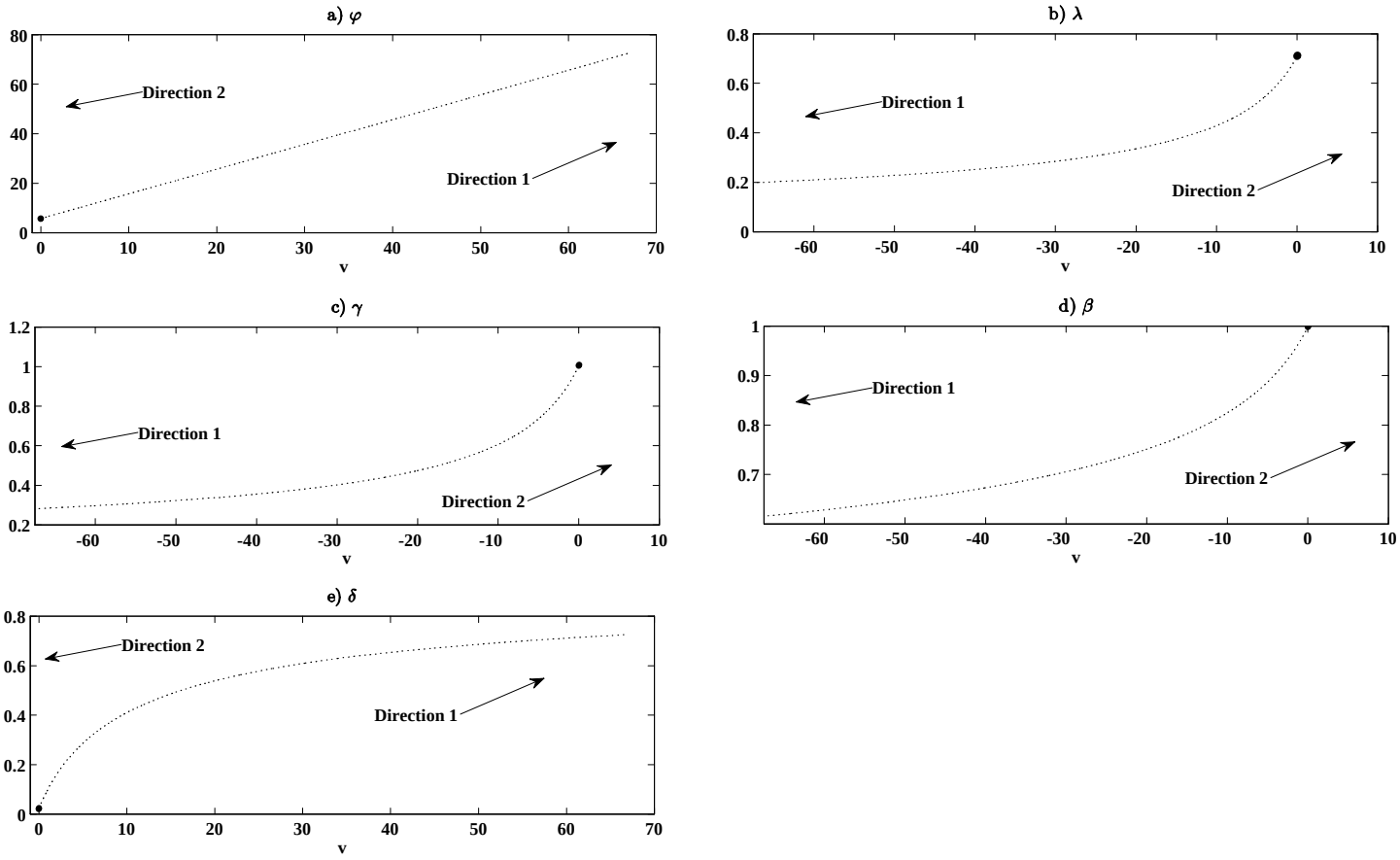
Note: 5% and 95% columns refer to the 5th and 95th percentiles of the distribution of the RWM draws.

Table 8: Log-likelihood and log-posterior values at posterior modes

	Posterior Mode			
	SW(2007)	Full Spectrum	Full Spectrum and Mean	BC Frequencies
Log-likelihood				
Full Spectrum	1186.91	**1212.25**	1211.88	543.27
Full Spectrum and Mean	1184.59	n/a	**1209.70**	n/a
BC Frequencies	254.00	258.09	257.81	**283.98**
	SW(2007)	Full Spectrum	Full Spectrum and Mean	BC Frequencies
Log-posterior				
Full Spectrum	1175.89	**1200.95**	1200.05	552.87
Full Spectrum and Mean	1172.05	n/a	**1197.38**	n/a
BC Frequencies	242.98	246.79	245.98	**293.58**

Note. Entries in the table correspond to the log-likelihoods/log-posteriors, as specified by row labels, evaluated at different posterior modes, which were computed by maximizing the log-posterior specified by column labels. For example, the upper left corner gives the value of the log-likelihood constructed using Fourier frequencies between $2\pi/T$ and $2\pi(T-1)/T$ with the parameter value set to the posterior mode of SW(2007).

Figure 1. The non-identification curve $(\varphi, \lambda, \gamma, \beta, \delta)$



Note. The non-identification curve is given by $\partial\theta(v)/\partial v = c(\theta),\ \theta(0) = \theta_0$, where $c(\theta)$ is the eigenvector corresponding to the only zero eigenvalue of $G(\theta)$. The approximation is computed recursively using the Euler method, so that $\theta(v_{j+1}) = \theta(v_j) + c(\theta(v_j))h,$ where h is the step size, fixed at 1e-04. $(\varphi, \lambda, \gamma, \beta, \delta)$ change simultaneously along the curve in the indicated directions. Directions 1 and 2 are obtained by restricting the first element of $c(\theta)$ to be positive or negative respectively. Since a discount rate greater than 1 contradicts economic theory, Direction 2 is truncated at the last point where β is below 1. The curve is extended for 670000 steps in Direction 1. Since there are only 472 steps in Direction 2, the respective curve appears as a bold dot on the graphs.

Figure 2(a). The estimated impulse responses of output to shocks

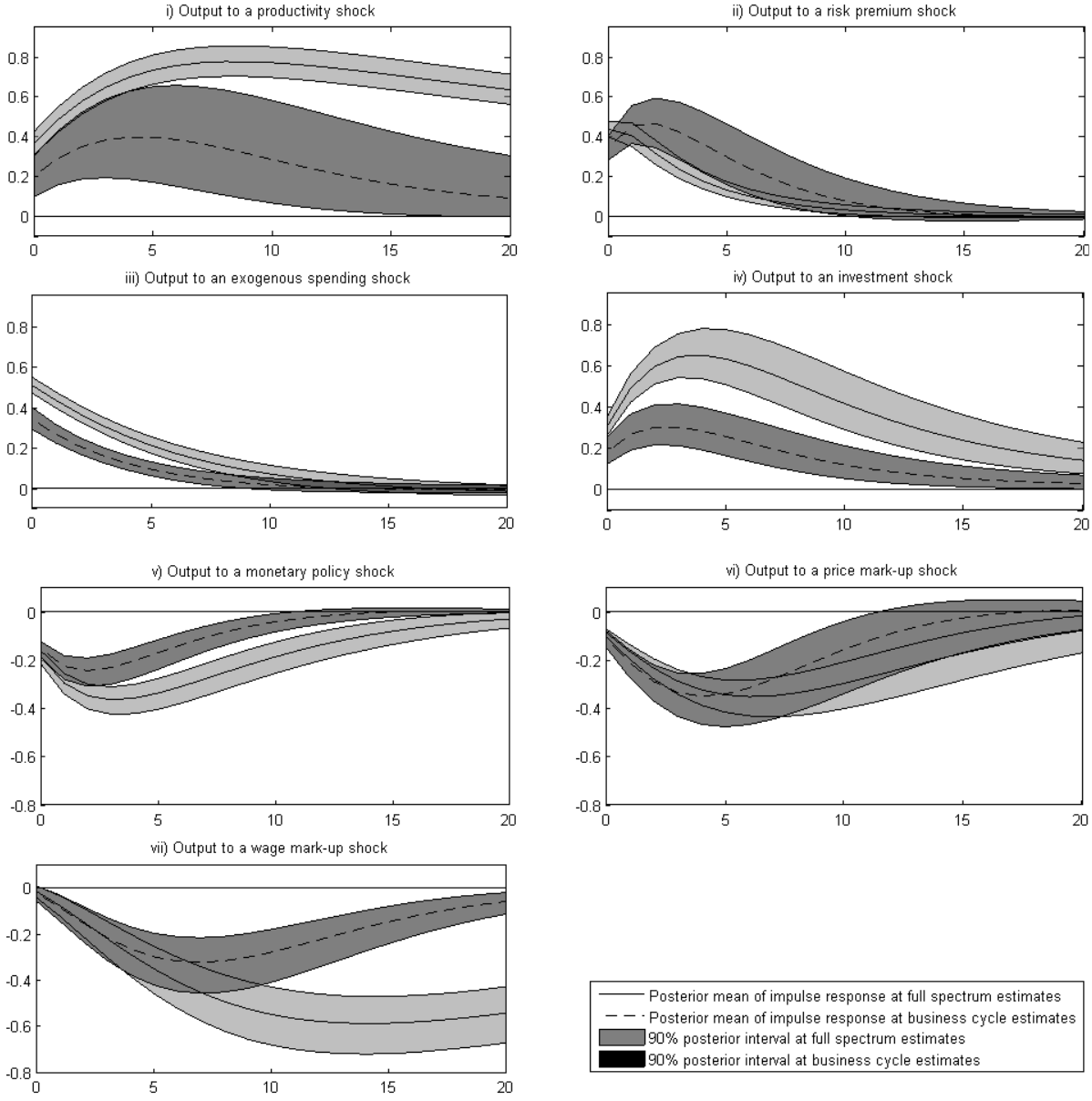


Figure 2(b). The estimated impulse responses of labor hours to shocks

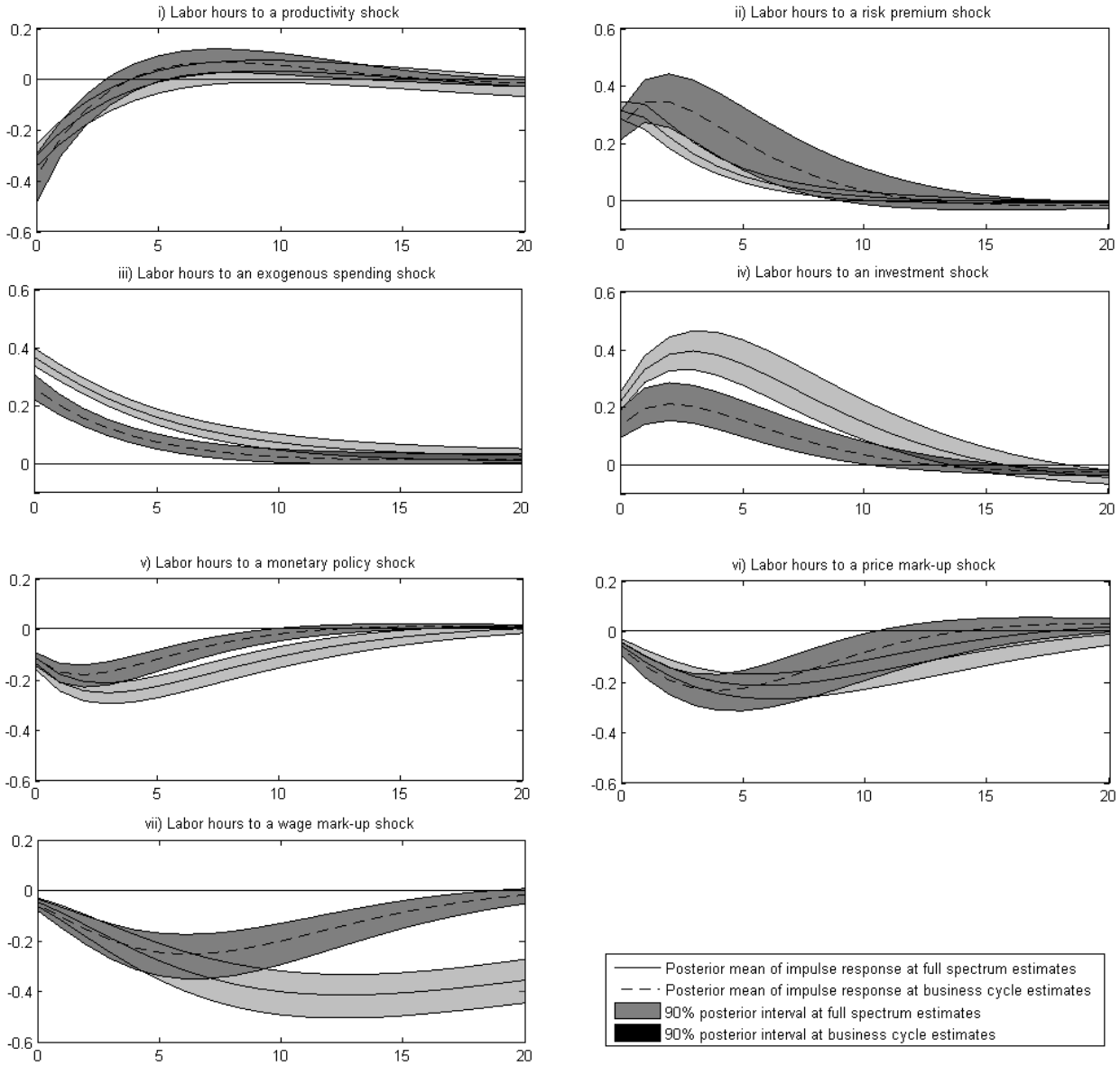


Figure 2(c). The estimated impulse responses of inflation to shocks

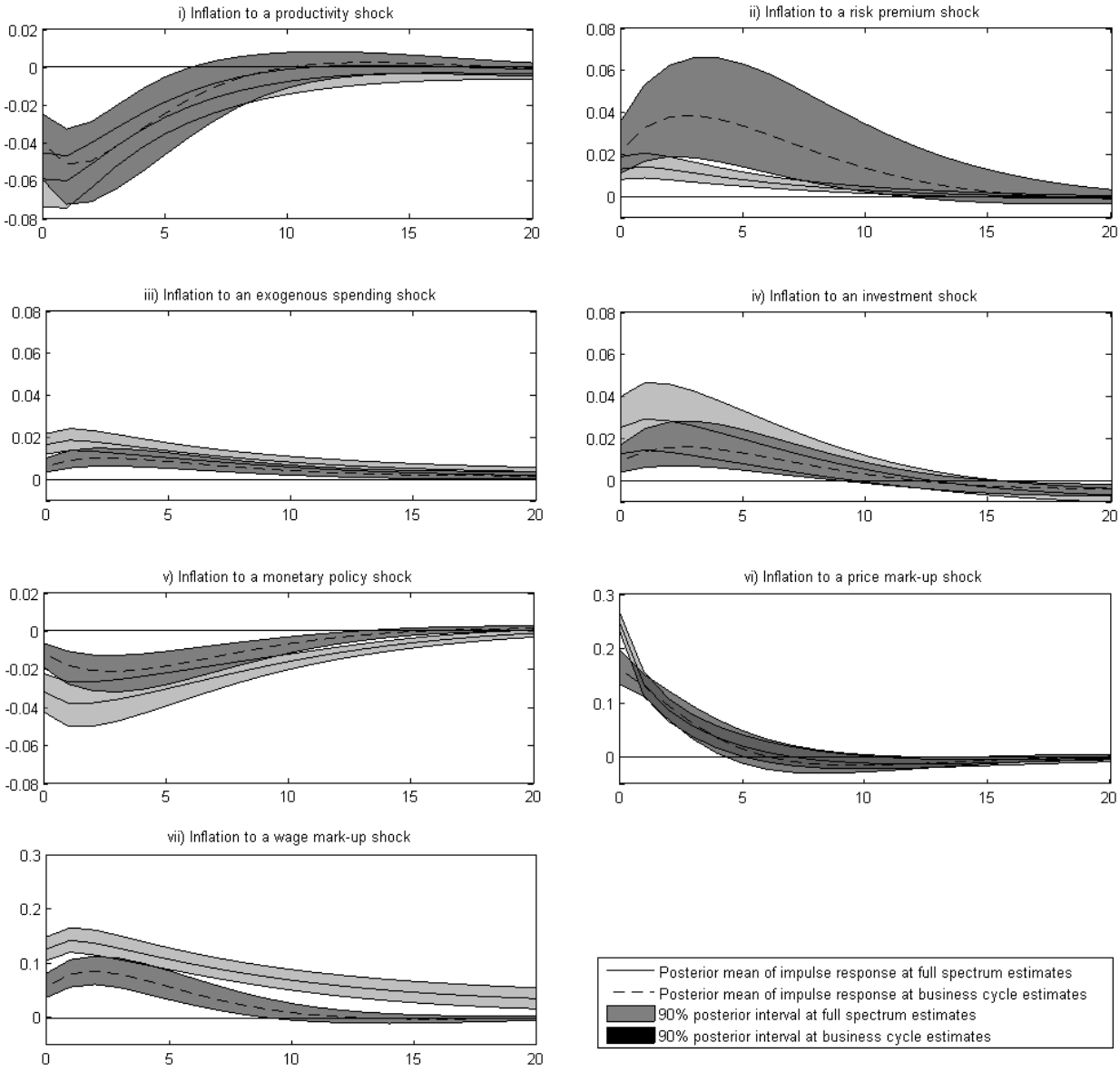


Figure 2(d). The estimated impulse responses of interest rate to shocks

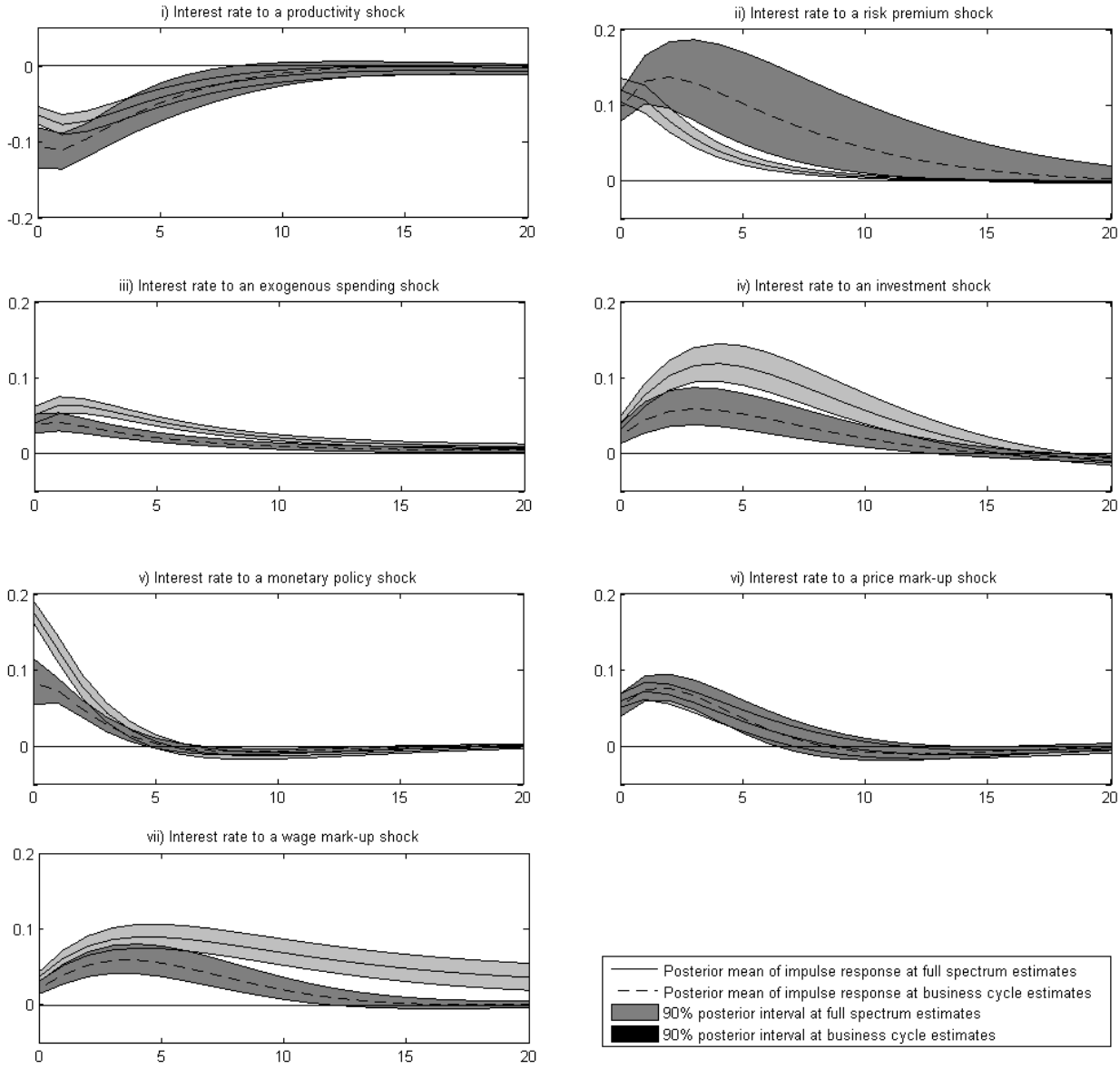


Figure 2(e). The estimated impulse responses of consumption to shocks

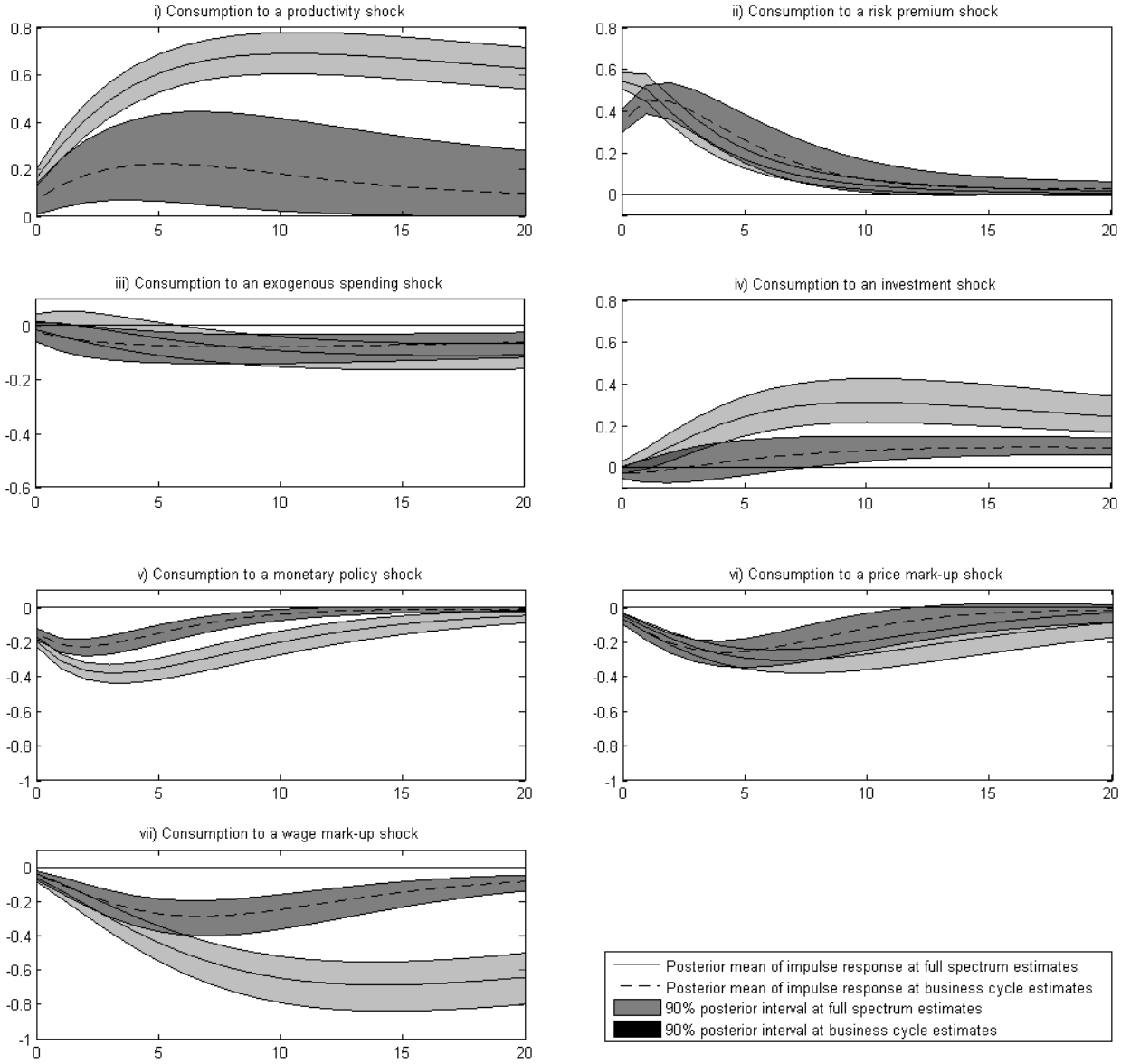


Figure 2(f). The estimated impulse responses of investment to shocks

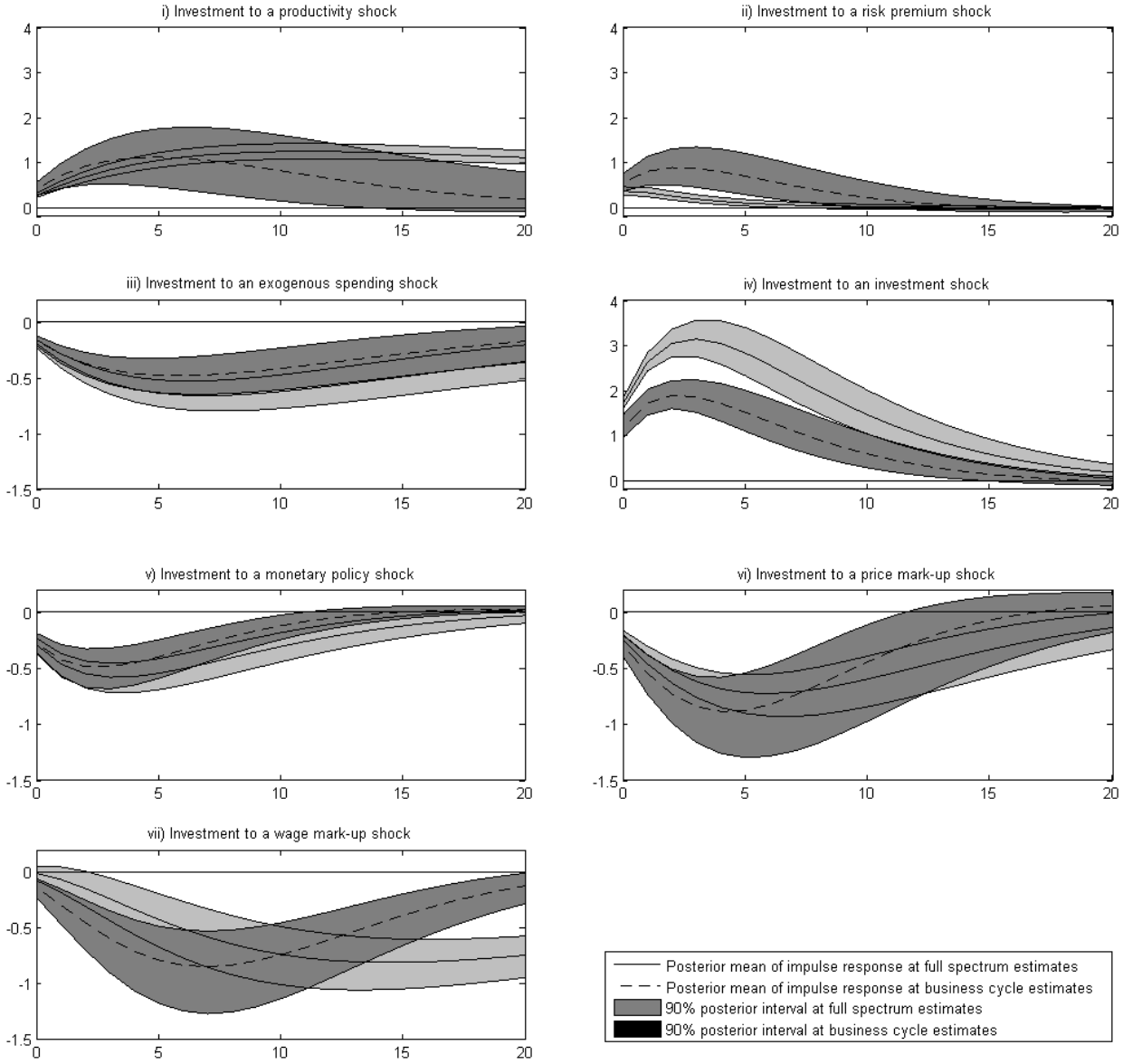


Figure 2(g). The estimated impulse responses of wage to shocks

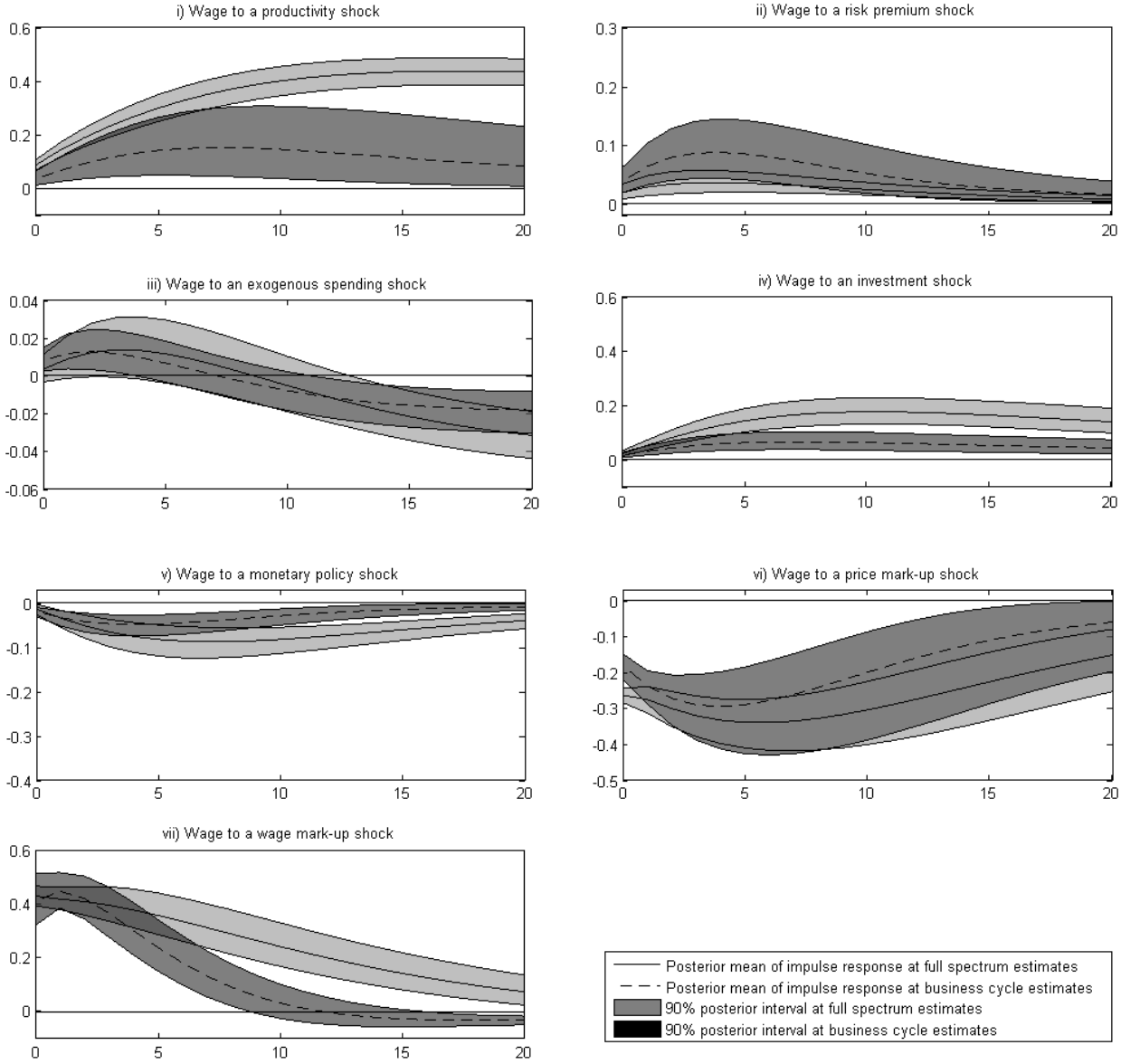


Figure 3(a). The estimated impulse responses of output to shocks

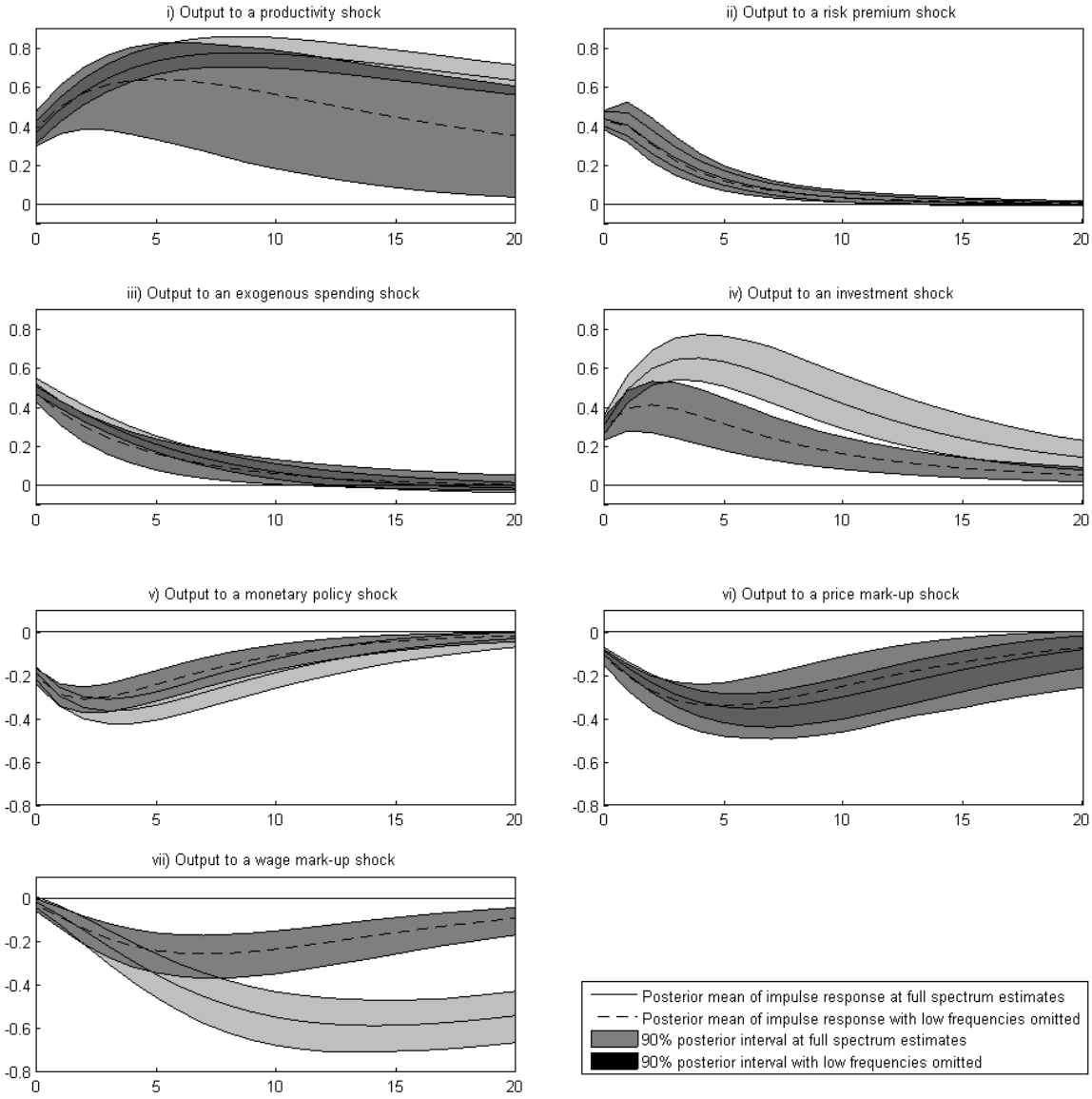


Figure 3(b). The estimated impulse responses of labor hours to shocks

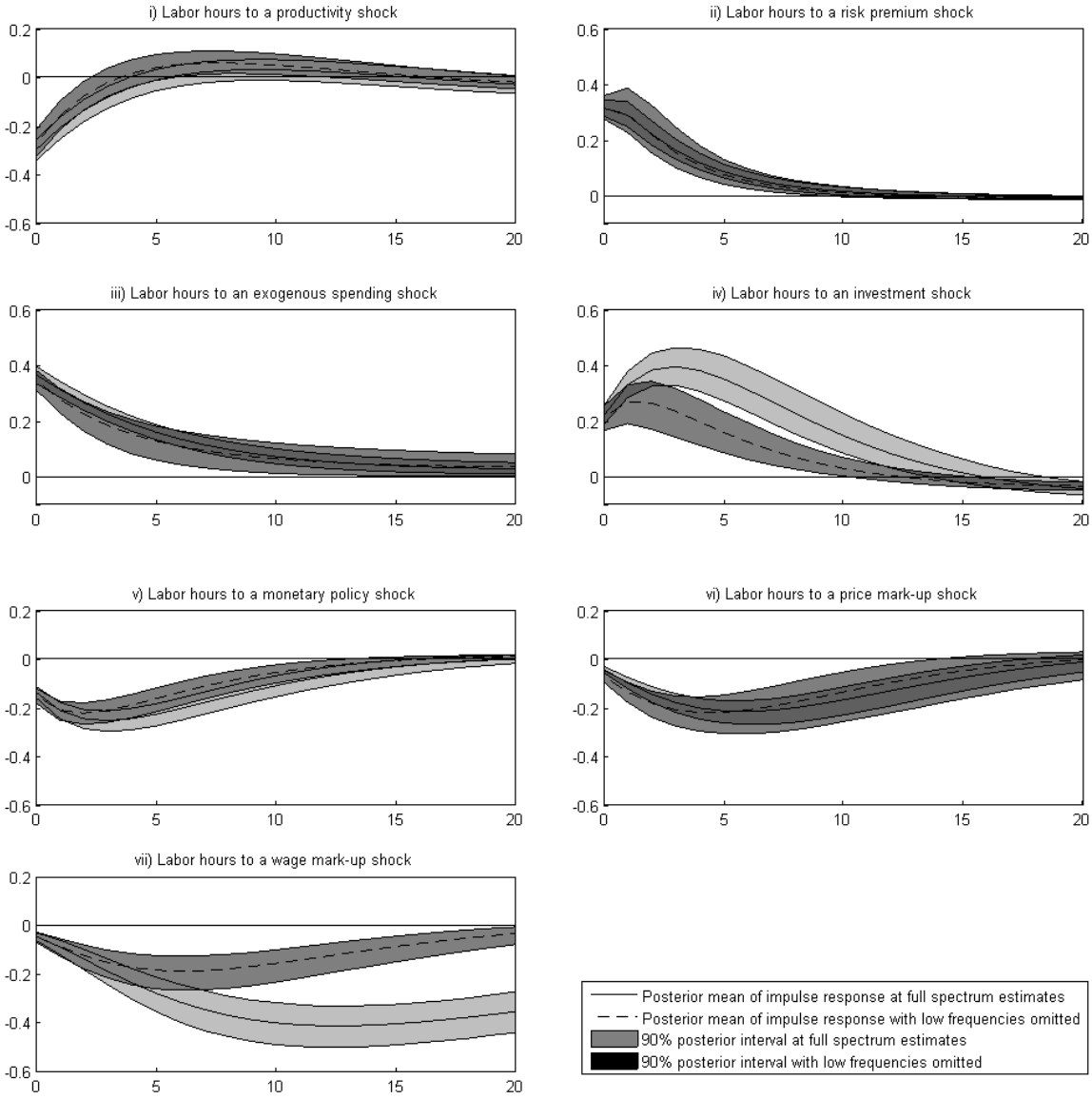


Figure 3(c). The estimated impulse responses of inflation to shocks

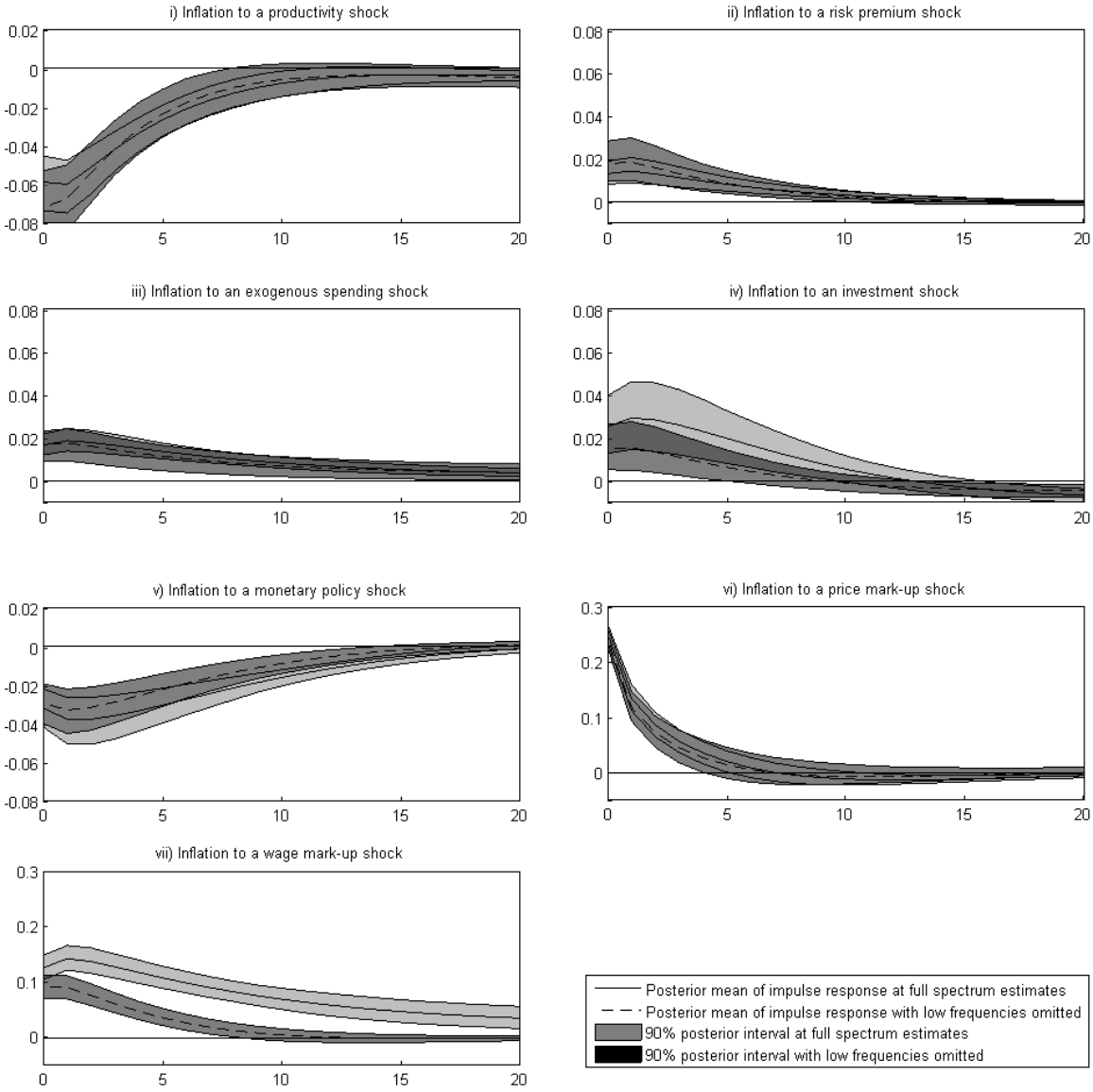


Figure 3(d). The estimated impulse responses of interest rate to shocks

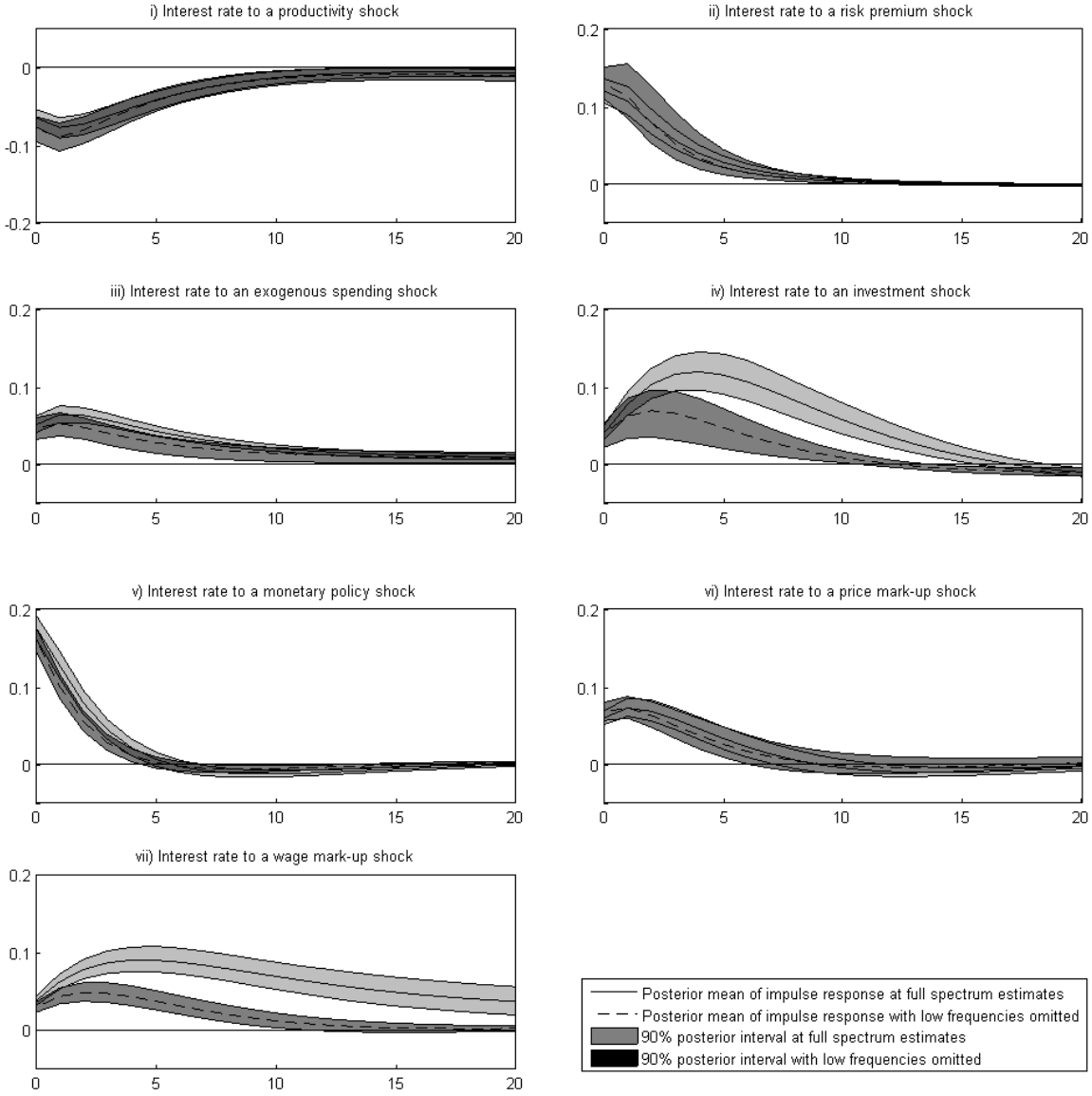


Figure 3(e). The estimated impulse responses of consumption to shocks

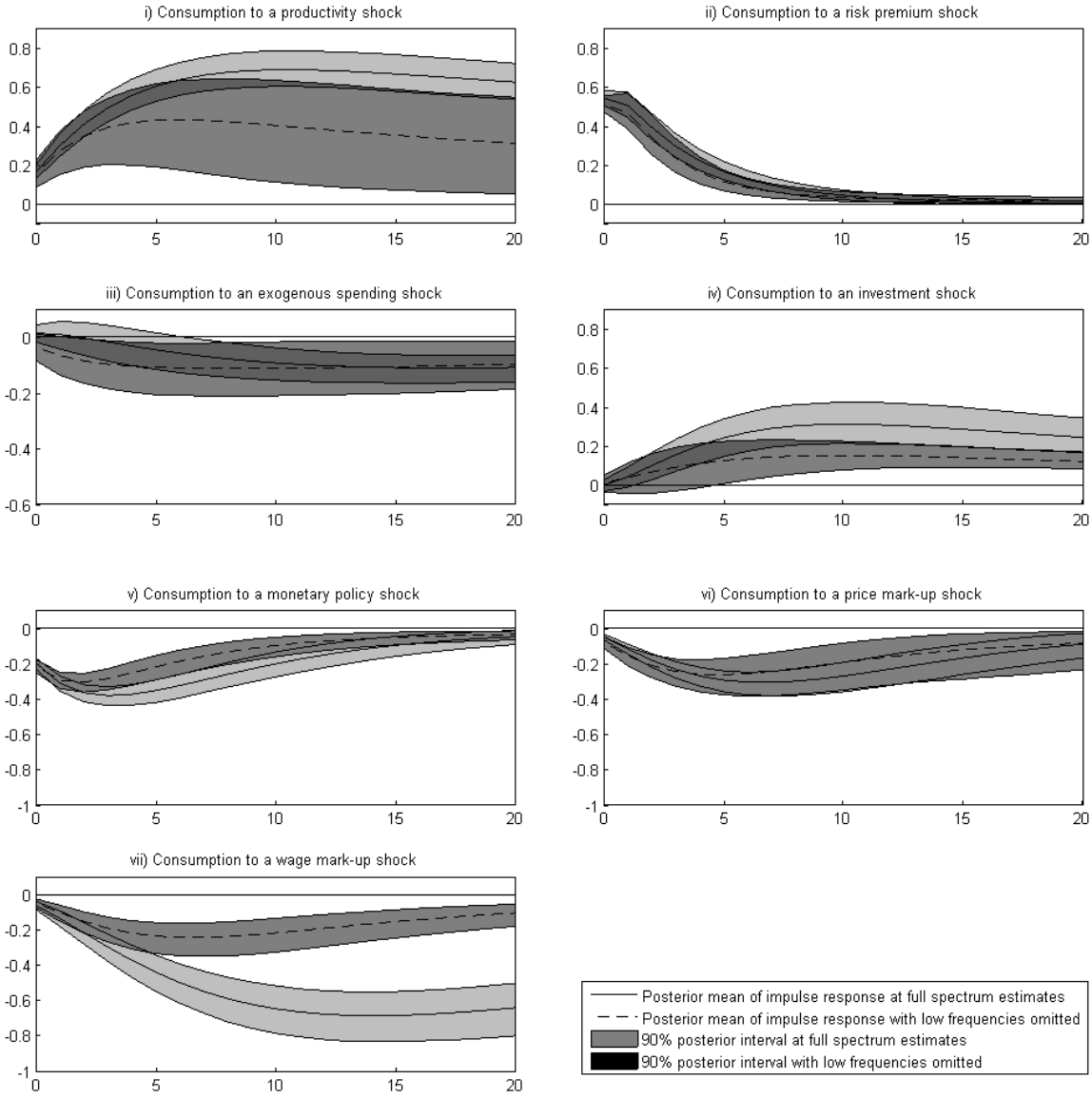


Figure 3(f). The estimated impulse responses of investment to shocks

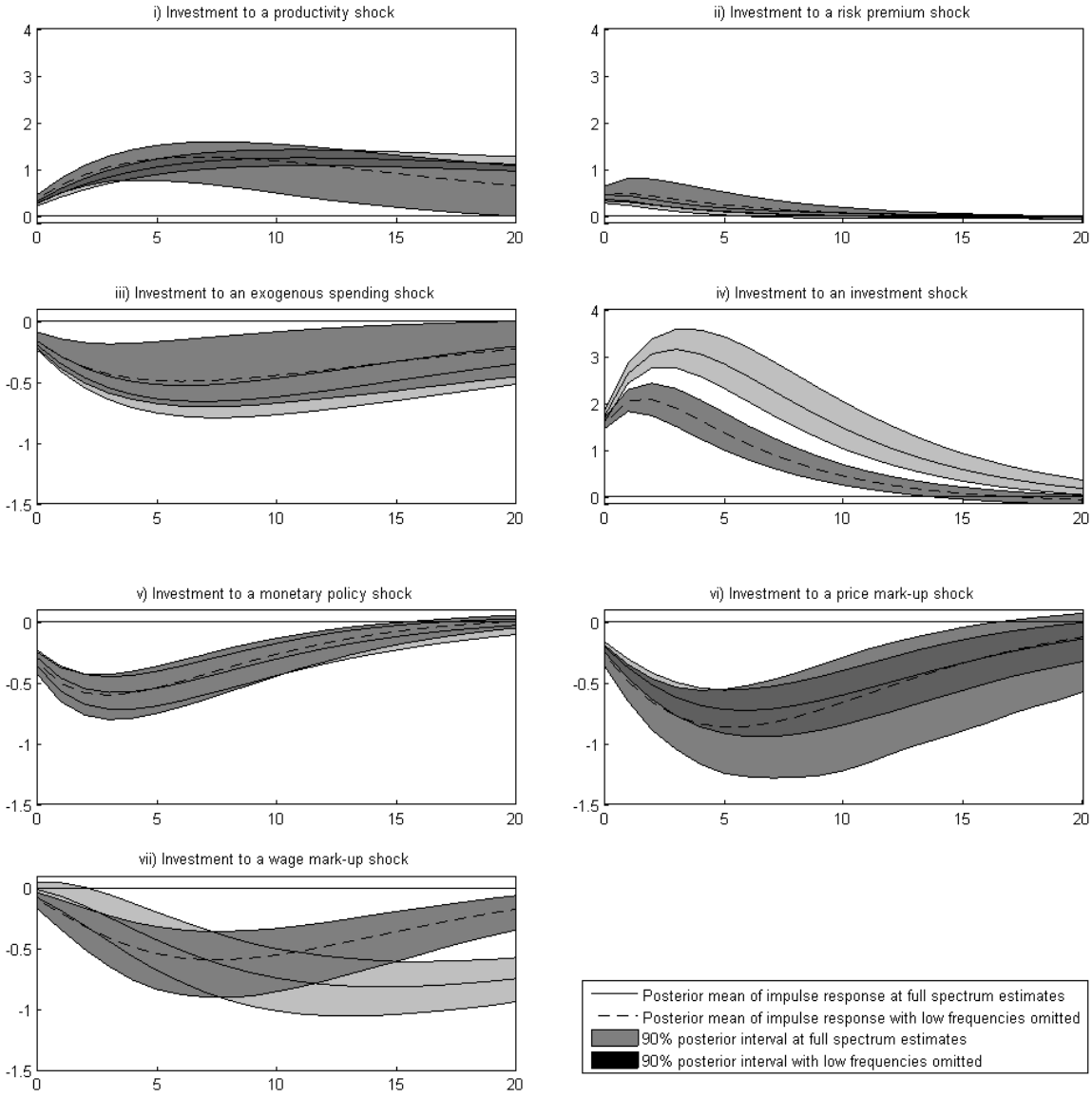


Figure 3(g). The estimated impulse responses of wage to shocks

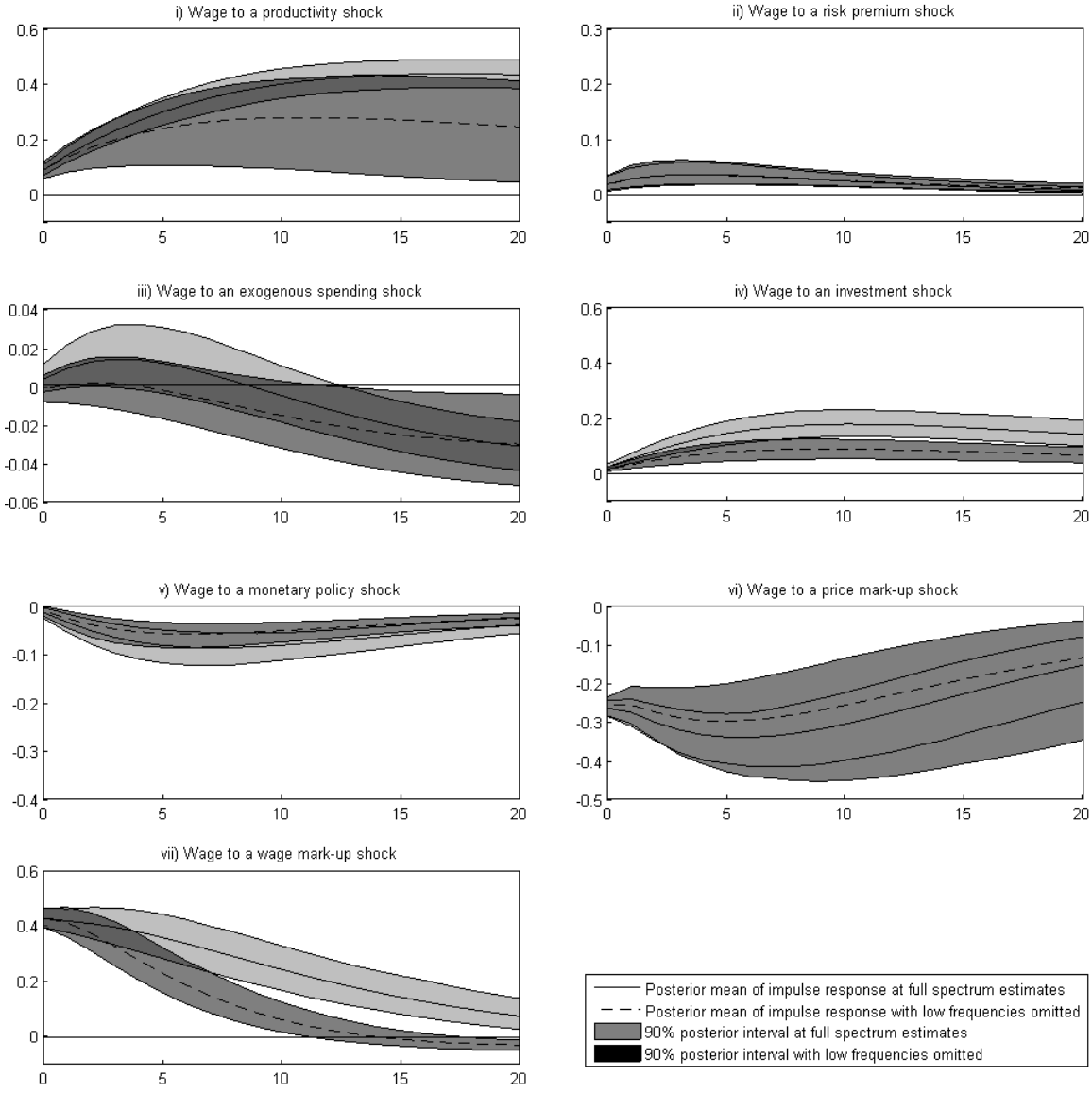


Figure 4(a). The estimated impulse responses of output to shocks

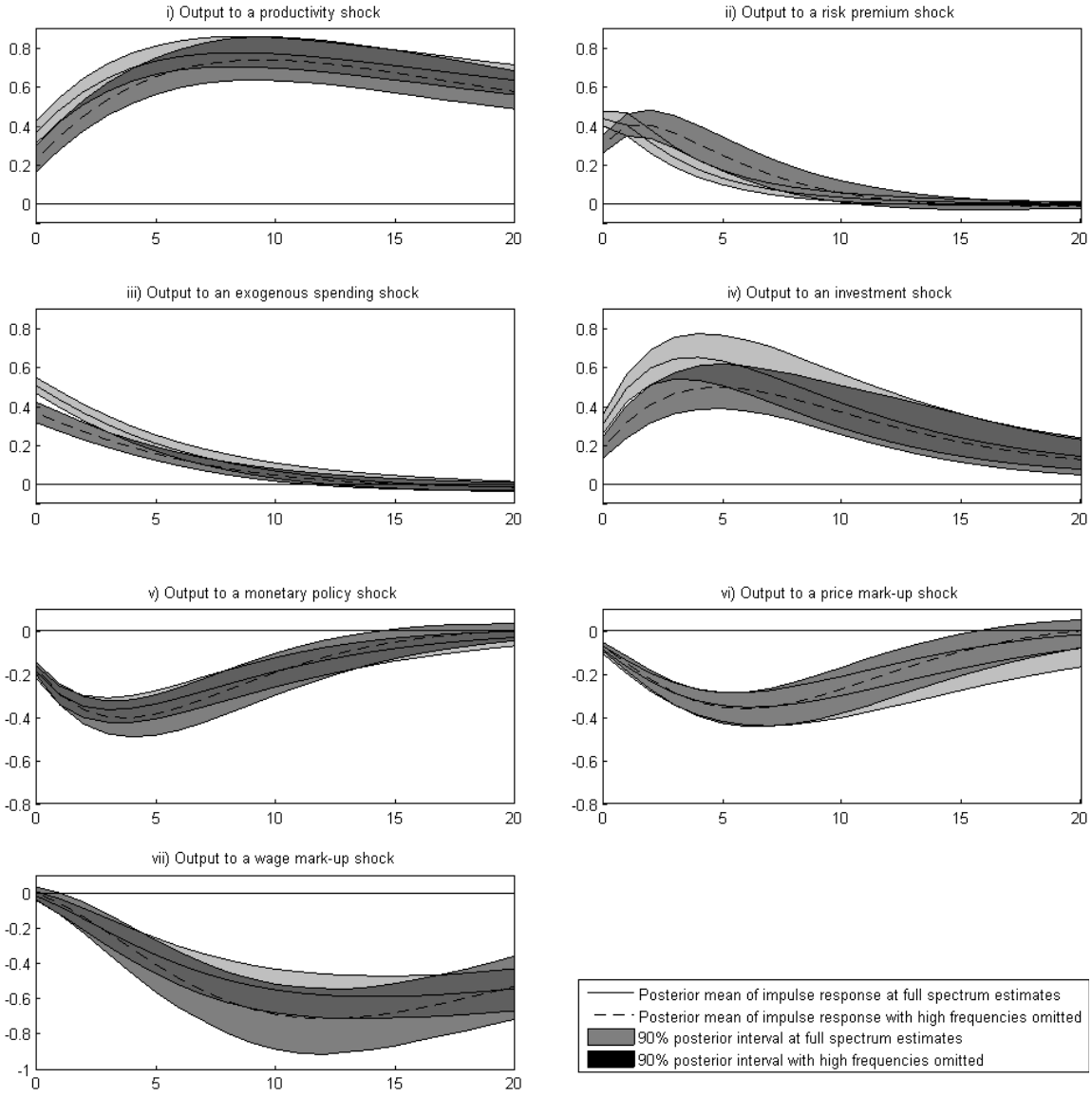


Figure 4(b). The estimated impulse responses of labor hours to shocks

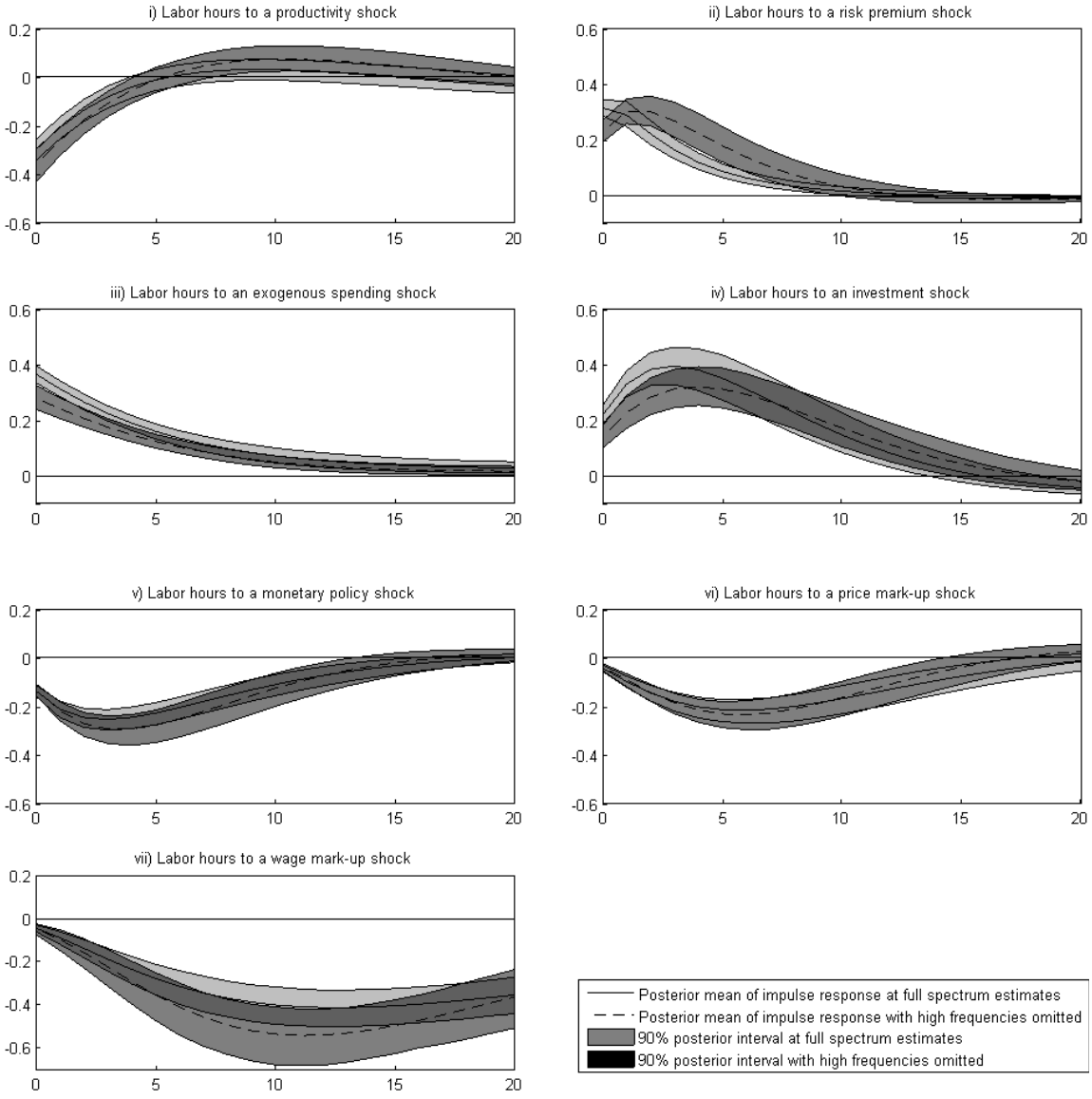


Figure 4(c). The estimated impulse responses of inflation to shocks

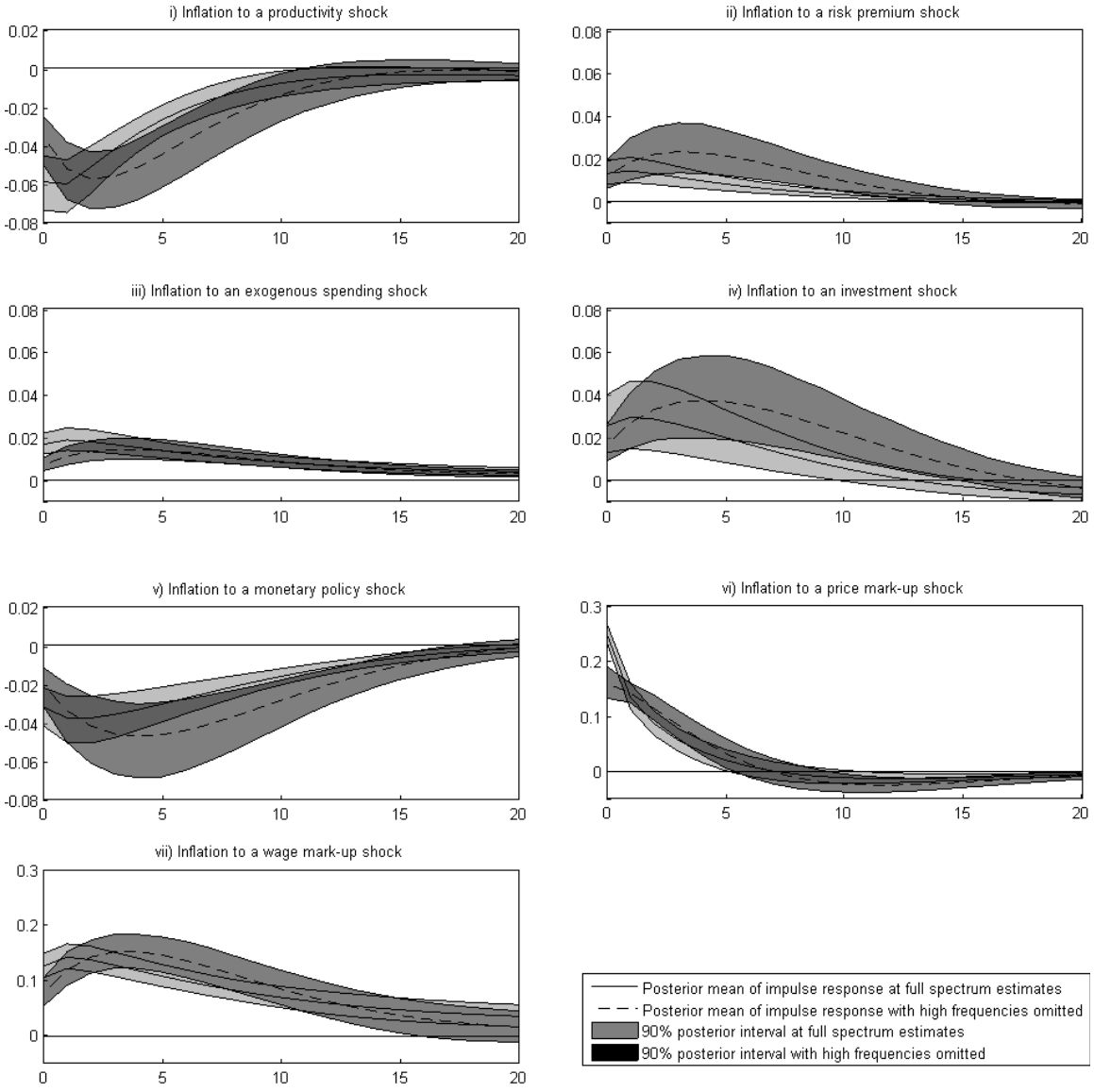


Figure 4(d). The estimated impulse responses of interest rate to shocks

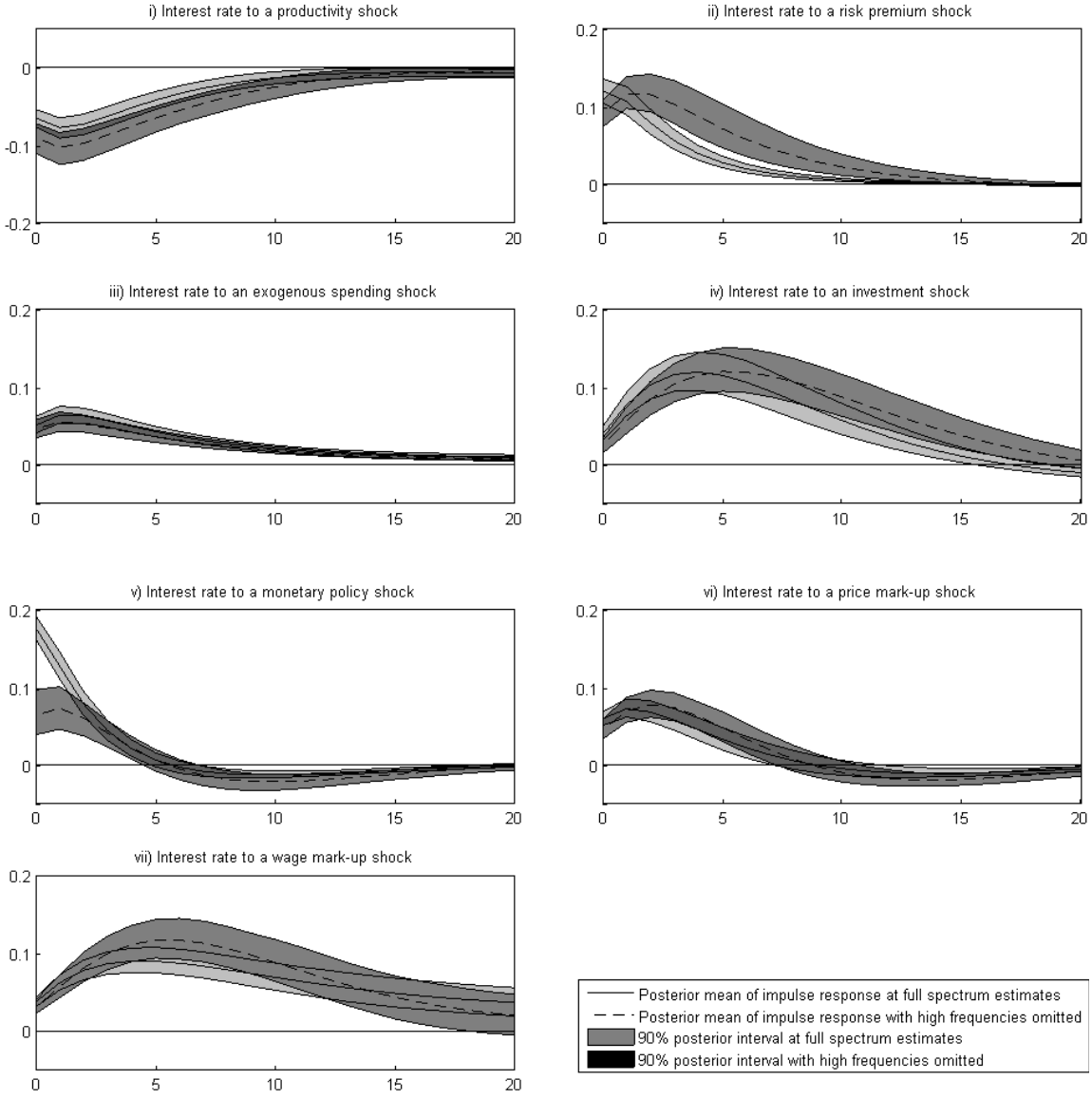


Figure 4(e). The estimated impulse responses of consumption to shocks

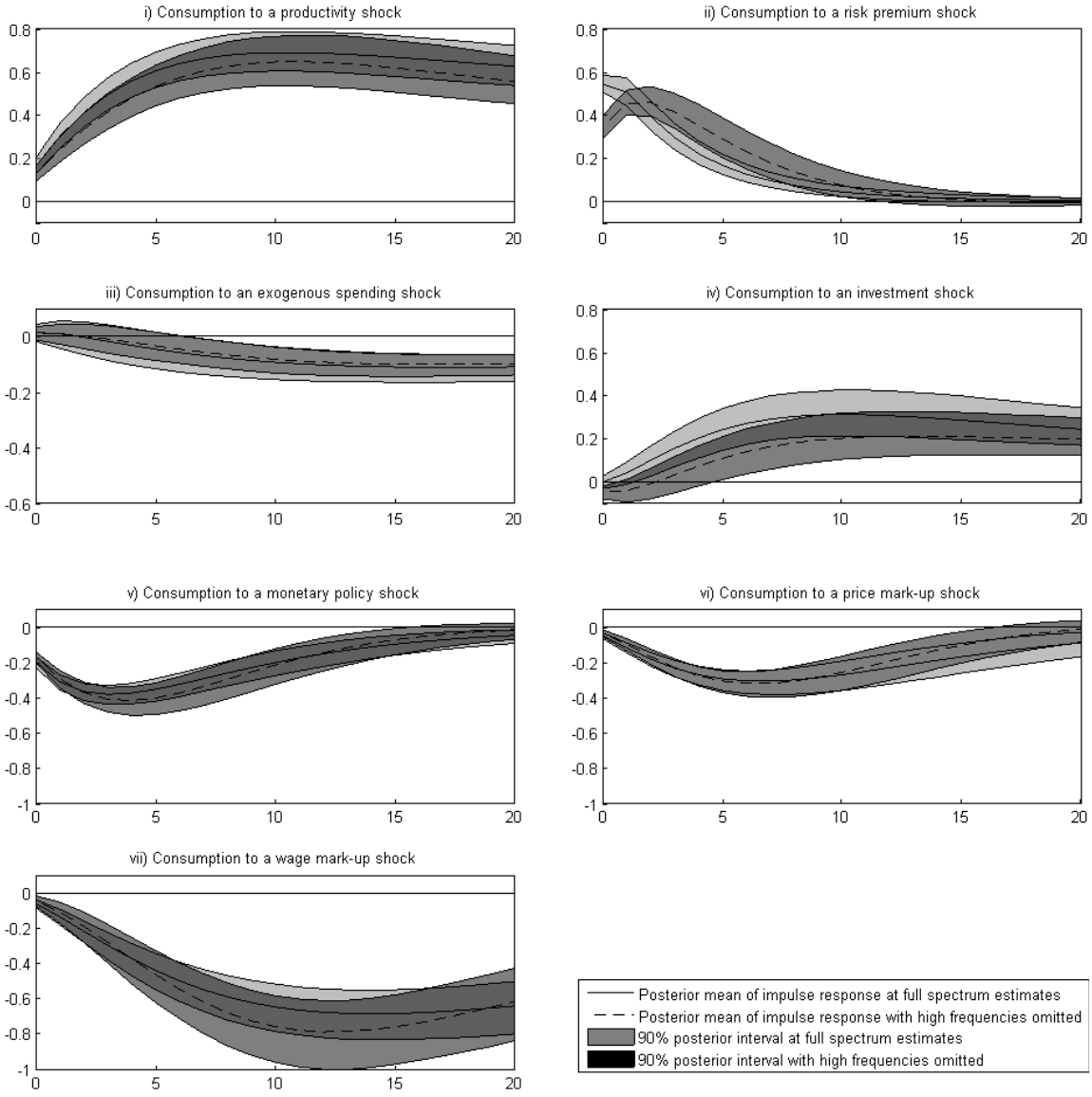


Figure 4(f). The estimated impulse responses of investment to shocks

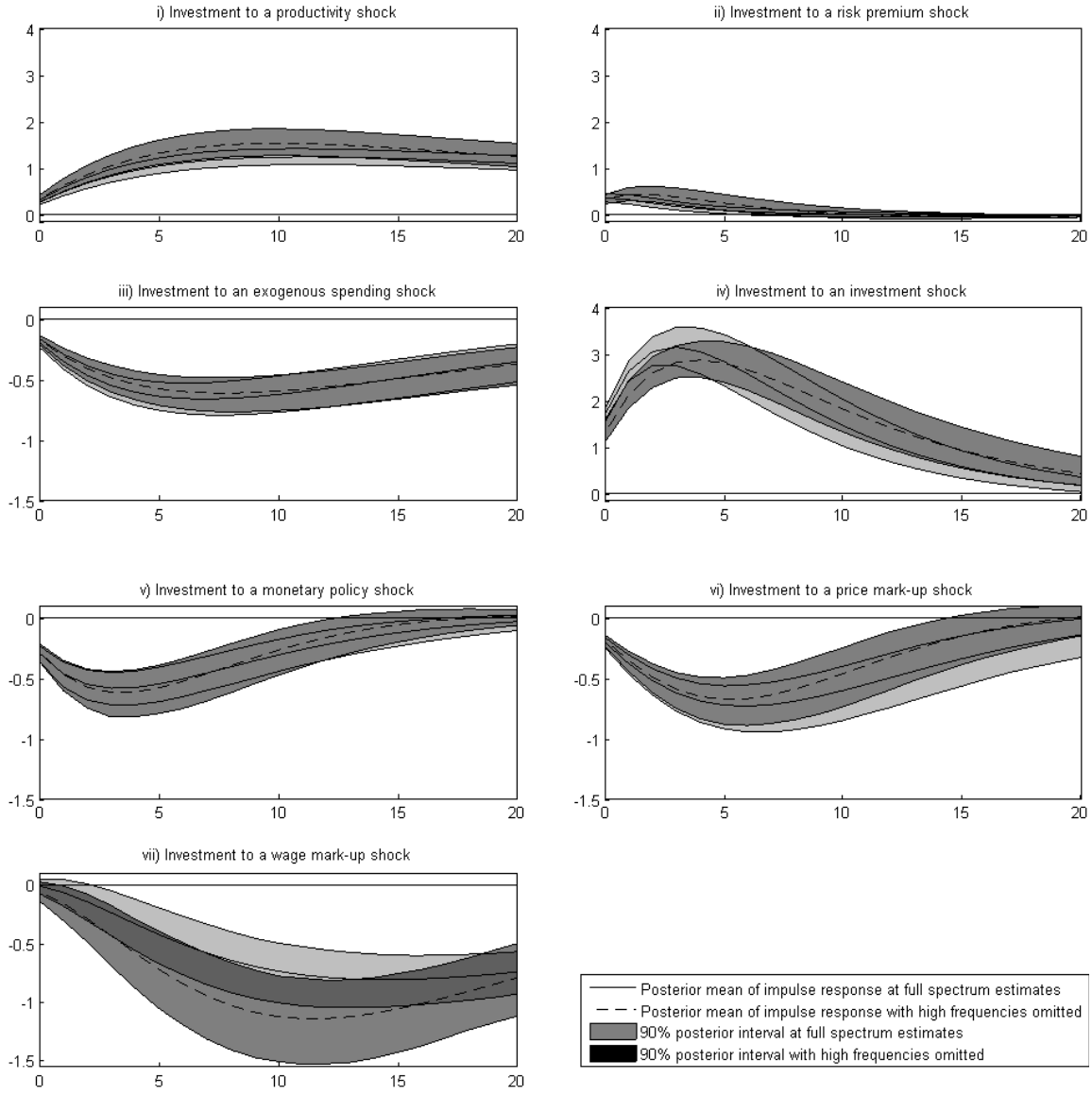


Figure 4(g). The estimated impulse responses of wage to shocks

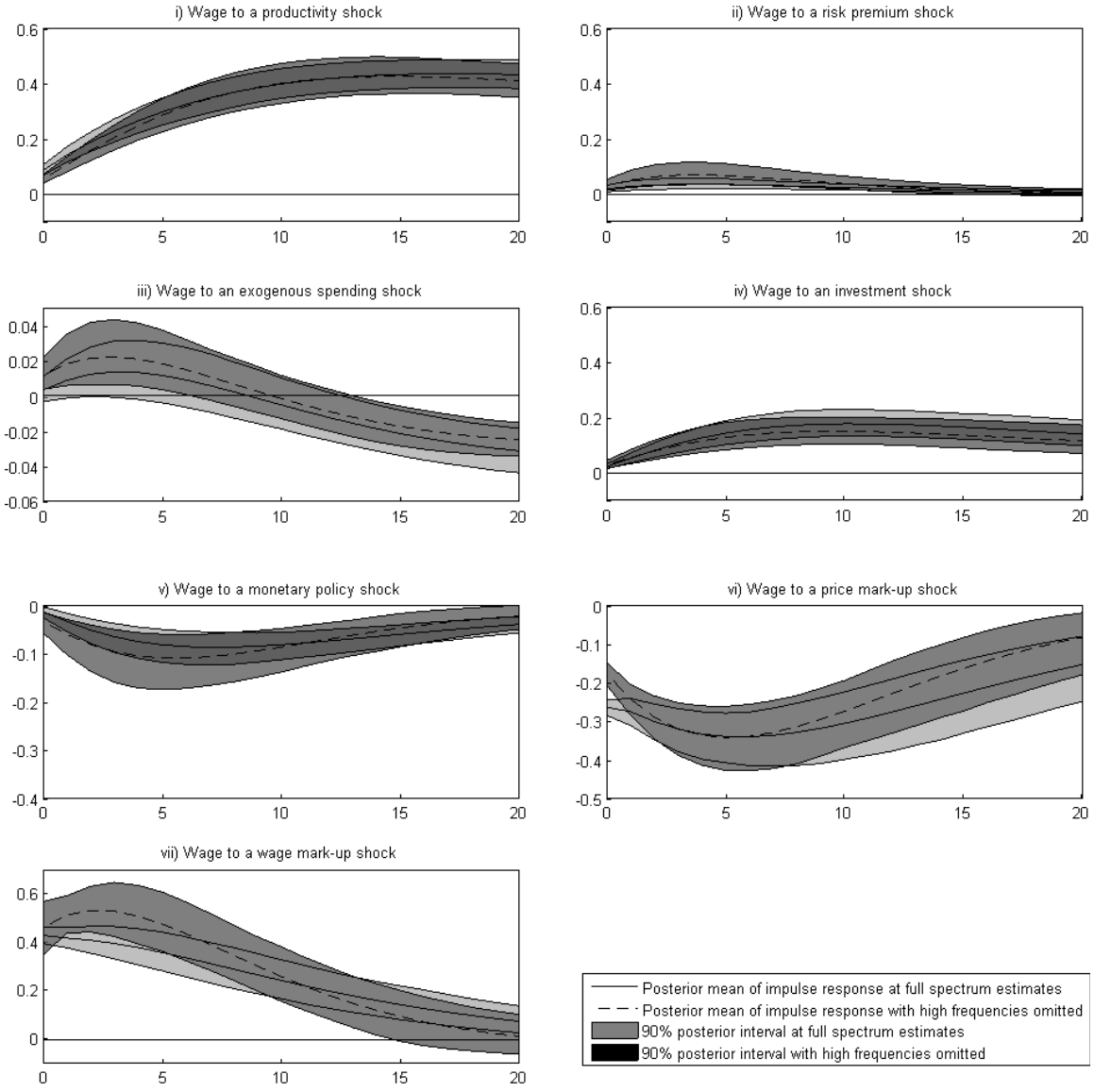


Figure 5. Model implied and nonparametrically estimated log spectra of observables

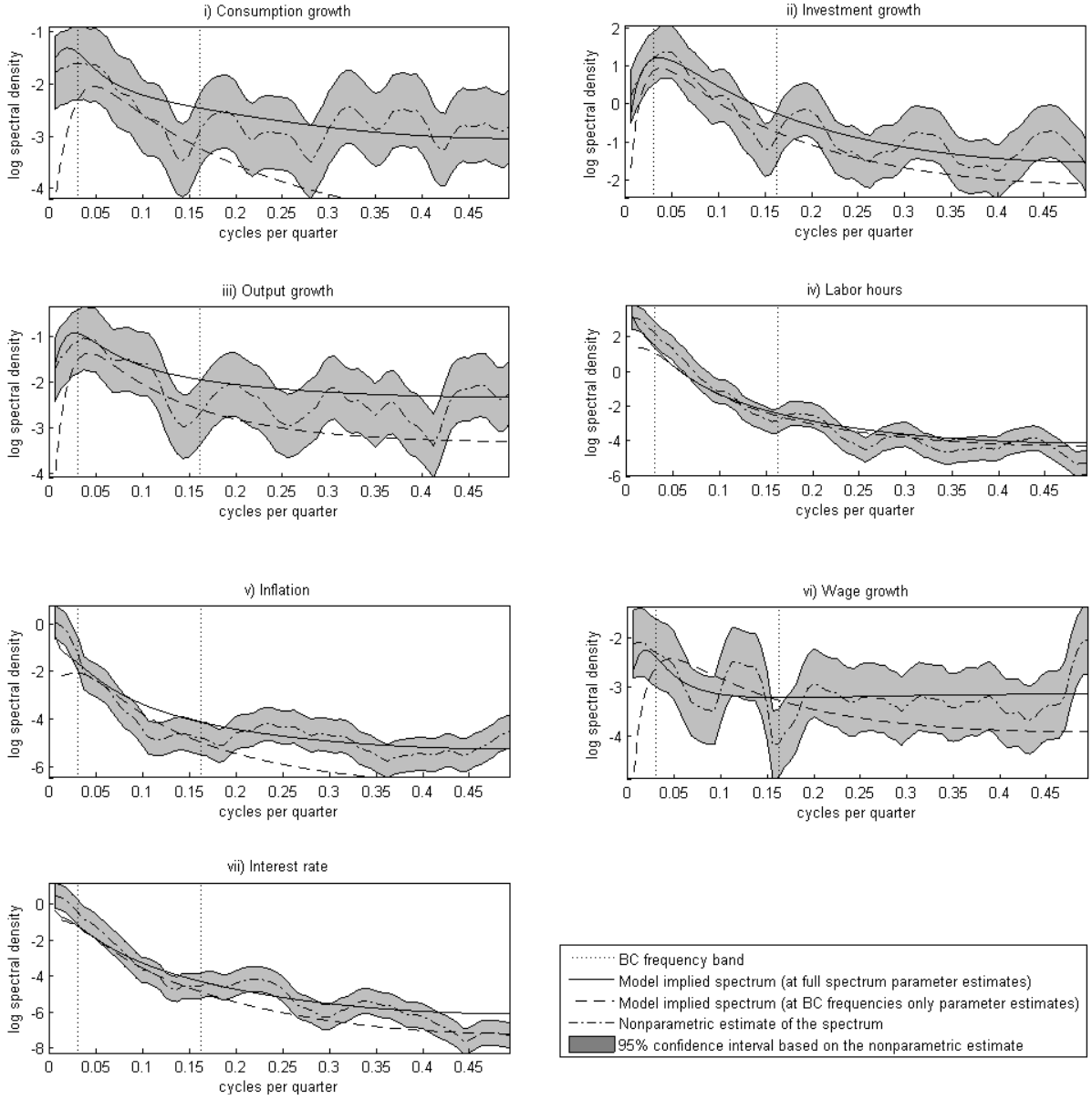


Figure 6(a). Model implied and nonparametrically estimated coherency between observables

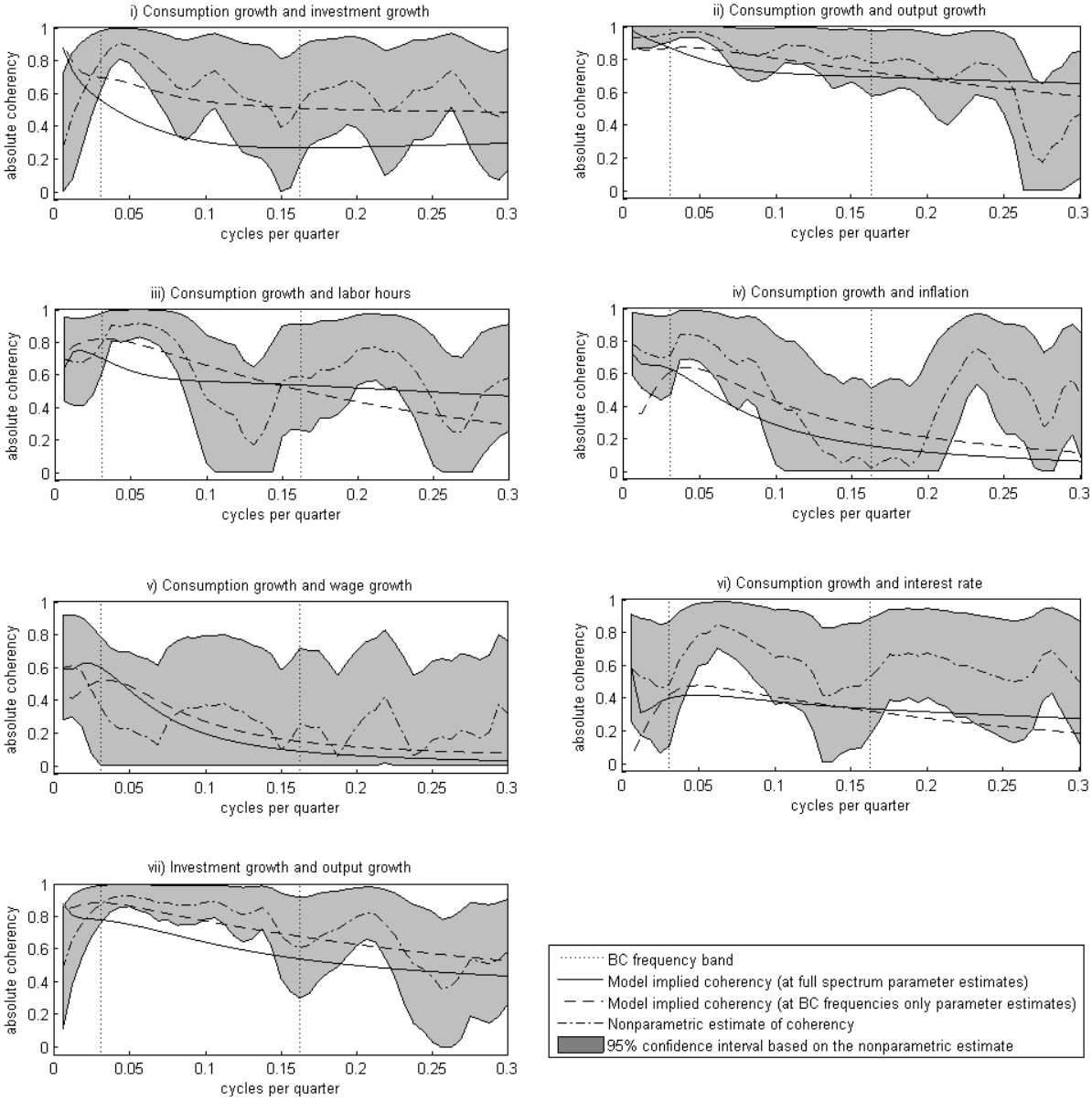


Figure 6(b). Model implied and nonparametrically estimated coherency between observables

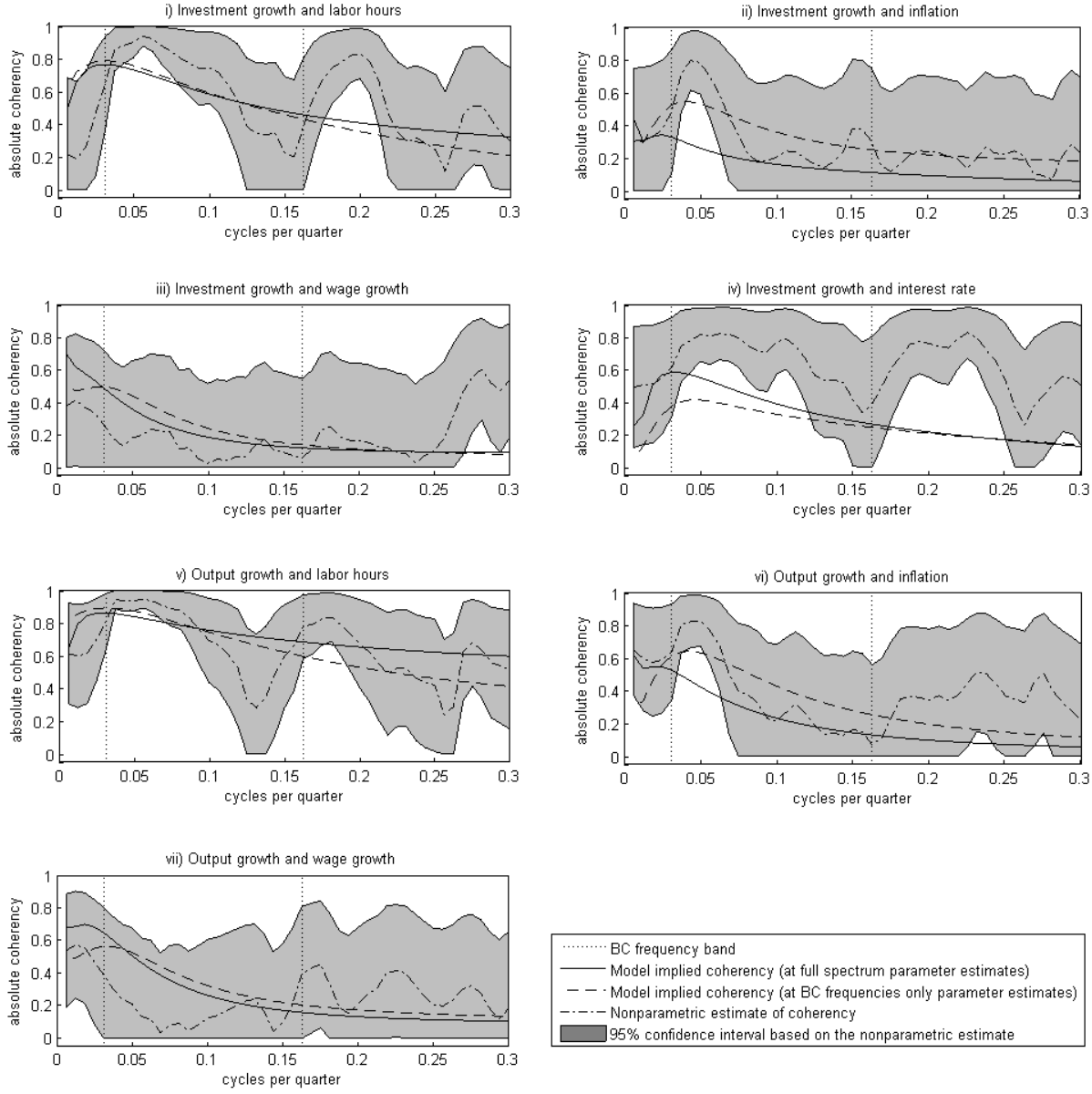


Figure 6(c). Model implied and nonparametrically estimated coherency between observables