

# Priors from Frequency-Domain Dummy Observations

Marco Del Negro

Francis X. Diebold

*Federal Reserve Bank of Atlanta*

*University of Pennsylvania*

Frank Schorfheide\*

*University of Pennsylvania*

February 12, 2008

*Very Preliminary and Incomplete*

## Abstract

By exploiting the insight that the misspecification of dynamic stochastic general equilibrium (DSGE) models is more prevalent at some frequencies than at others, we develop methods that enable different degrees of relaxation of the DSGE restrictions in different directions. We approximate the DSGE model by a vector autoregression. Dummy observations are constructed from the DSGE model and converted into the frequency domain. By re-weighting the frequency domain dummy observations we can control the extent to which the restrictions derived from economic theory are relaxed. Bayesian marginal data densities can then be used to obtain a data-driven procedure that determines the optimal degree of shrinkage toward the DSGE model restrictions. We provide several numerical illustrations of our procedure.

JEL CLASSIFICATION: C32, E52, F41

KEY WORDS: Bayesian Econometrics, DSGE Models,  
Frequency Domain Analysis, Misspecification

---

\*We thank Sungbae An for his excellent research assistance.

## 1 Introduction

This paper exploits the insight that the misspecification of dynamic stochastic general equilibrium (DSGE) models is more prevalent at some frequencies than at others, developing methods that enable different degrees of relaxation of the DSGE restrictions in various directions. For example, DSGE models impose very strong long-run restrictions. In the neoclassical growth model with a random walk technology process, output, consumption, investment, real wages, and the capital stock share a common stochastic trend, implying that pairwise ratios of those variables should be stationary (see King, Plosser, Rebelo, 1998), but a close look at the data suggests otherwise. Data-based violation of those long-run restrictions results in poor DSGE model fit, in particular compared to VARs that allow for more general common trend features. DSGE models, however, are designed for business cycle analysis; that is, they are designed to explain medium-term business cycle fluctuations, not very long-run or very short-run fluctuations. Hence we are much more willing to relax the very short-run and very long-run DSGE model restrictions than the more relevant medium-run DSGE model restrictions. Unfortunately, standard procedures do not permit this. The methods proposed in this paper do.

Del Negro and Schorfheide (2004) developed a framework in which a DSGE model was used to derive restrictions for vector autoregressions (VAR). Rather than imposing these restrictions dogmatically, Del Negro and Schorfheide constructed a family of prior distributions that concentrates much of its probability mass in the neighborhood of these restrictions. The prior has the property that it biases the VAR coefficient estimates toward the restrictions implied by a fully-specified dynamic model. Loosely speaking, the prior is implemented by augmenting the actual observations by dummy observations generated from the DSGE model, very much in the spirit of the classic Theil-Goldberger (1961) mixed estimation. The more of these dummy observations are added, the closer the VAR estimates stay to the DSGE model restrictions. This so-called DSGE-VAR framework can be used to estimate DSGE and VAR parameters, to evaluate DSGE model, and to forecast and conduct policy analysis, e.g., Del Negro and Schorfheide (2005), Del Negro, Schorfheide, Smets, and Wouters (2006), and Adolfson, Laséen, Lindé, and Villani (2006).

In this paper we extend the DSGE-VAR framework by considering dummy observations

from a DSGE model that have been transformed into the frequency domain and re-weighted to emphasize certain spectral bands along which the DSGE model fits well. The paper is organized as follows. Section 2 provides some evidence that the current generation of DSGE models is severely misspecified in terms of their low frequency implications. We consider a stochastic growth model with a number of frictions that include capital and labor adjustment costs. This model is essentially a flexible price and wages version of the medium-scale DSGE models that are currently used for applied monetary policy analysis, e.g., Smets and Wouters (2003). We document that this model is unable to generate the persistence in the *great ratios*, in particular the consumption-output ratio, that we observe in quarterly U.S. data. Section 3 briefly reviews the time-domain DSGE-VAR framework. Frequency-domain dummy observations are introduced in Section 4, Section 5 contains two illustrative examples, and Section 6 an (currently incomplete) empirical application. We conclude in Section 7 and outline future research.

## 2 Common Trends in U.S. Data and an Estimated DSGE Model

To illustrate that model misspecification may be more prevalent at some frequencies than at others we use a one-sector neoclassical growth model with several real frictions, based on work of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2003), including capital and labor adjustment costs. We abstract from nominal rigidities. Technology shifts according to an integrated labor augmenting exogenous process that induces a stochastic growth path along which output, consumption, and investment grow at the same rate and hours worked is stationary. We compute prior and posterior predictive densities for the spectrum of some of the great ratios (Klein and Kosobud, 1961) and compare them to spectral estimates constructed from actual U.S. data. Before presenting the empirical results we briefly outline the DSGE model.

## 2.1 The DSGE Model

A representative household maximizes the expected discounted lifetime utility from consumption  $C_t$  and hours worked  $L_t$ : given by:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s \left[ \log(C_{t+s} - hC_{t+s-1}) - \frac{\phi_{t+s}}{1 + \nu_l} L_{t+s}^{1+\nu_l} \right]. \quad (1)$$

Household's preferences display habit-persistence. The short-run (Frisch) labor supply elasticity is  $\nu_l$ . The exogenous process

$$\ln \phi_t = (1 - \rho_\phi) \ln \phi + \rho_\phi \ln \phi_{t-1} + \sigma_\phi \epsilon_{\phi,t}$$

can be interpreted as labor supply shock, since an increase of  $\phi_t$  raises aggregate labor supply. This may reflect permanent shifts in per capita hours of work due to demographic changes, tax reforms, shifts in the marginal rate of substitution between leisure and consumption, or (non-neutral) technological changes in household production technology.

The household supplies labor at the competitive equilibrium wage  $W_t$  and rents capital services to the firms at the competitive rental rate  $R_t^k$ . The household's budget constraint is given by:

$$C_{t+s} + I_{t+s} + T_{t+s} \leq A_{t+s-1} + \Pi_{t+s} + W_{t+s} L_{t+s} + [R_{t+s}^k u_{t+s} \bar{K}_{t+s-1} - a(u_{t+s}) \bar{K}_{t+s-1}], \quad (2)$$

where  $I_t$  is investment,  $\Pi_t$  is the profit the household gets from owning firms,  $W_t$  is the real wage earned by the household, and  $T_t$  are lump-sum taxes (transfers) from the government. The term within parenthesis represents the return to owning  $\bar{K}_t$  units of capital. Households choose the utilization rate of their own capital,  $u_t$ . Households rent to firms in period  $t$  an amount of effective capital equal to:

$$K_t = u_t \bar{K}_{t-1}, \quad (3)$$

and receive  $R_t^k u_t \bar{K}_{t-1}$  in return. They however have to pay a cost of utilization in terms of the consumption good equal to  $a(u_t) \bar{K}_{t-1}$ . Households accumulate capital according to the equation:

$$\bar{K}_t = (1 - \delta) \bar{K}_{t-1} + \mu_t \left( 1 - S \left( \frac{I_t}{I_{t-1}} \right) \right) I_t, \quad (4)$$

where  $\delta$  is the rate of depreciation, and  $S(\cdot)$  is the cost of adjusting investment, with  $S(e^\gamma) = 0$ , and  $S''(\cdot) > 0$ . The term  $\mu_t$  is a stochastic disturbance to the price of investment

relative to consumption, see Greenwood, Hercowitz, and Krusell (1998), which follows the exogenous process:

$$\ln \mu_t = (1 - \rho_\mu) \ln \mu + \rho_\mu \ln \mu_{t-1} + \sigma_\mu \epsilon_{\mu,t}. \quad (5)$$

Firms rent capital, hire labor and capital services, and produce final goods according to the following technology

$$Y_t = (Z_t L_t)^{1-\alpha} K_t^\alpha \left( 1 - \varphi \cdot \left( \frac{L_t}{L_{t-1}} - 1 \right)^2 \right), \quad (6)$$

where the technology shock  $Z_t$  (common across all firms) follows a unit root process in logs:

$$z_t = \ln(Z_t/Z_{t-1}) = \gamma + \sigma_z \epsilon_{z,t}. \quad (7)$$

The last term in (6) captures the cost of adjusting labor inputs:  $\varphi \geq 0$ . In models  $\mathcal{M}_0$  and  $\mathcal{M}_1$ , there is no adjustment cost:  $\varphi = 0$ . Despite various types of adjustment costs in the labor market – e.g., search (Andolfatto, 1996), learning (Chang, Gomes, and Schorfheide, 2002), time non-separable utility in leisure (Kydland and Prescott, 1982) – we use a simple reduced-form quadratic cost to firms without taking a particular stand on the micro foundations of the nature of friction. The firms maximize expected discounted future profits

$$\mathbb{E}_t \left[ \sum_{s=0}^{\infty} \beta^{t+s} \Xi_{t+s|t} \Pi_t \right], \quad (8)$$

where  $\Pi_t = Y_t - W_t L_t - R_t^k K_t$  and  $\Xi_{t+s|t}$  is the marginal value of a unit consumption to a household, which is treated as exogenous to the firm.

A fraction of aggregate output is purchased by the government:

$$G_t = (1 - 1/g_t) Y_t, \quad (9)$$

where  $g_t$  follows the exogenous process:

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \sigma_g \epsilon_{g,t} \quad (10)$$

The government levies lump-sum taxes  $T_t$  to finance its purchases. In equilibrium the goods, labor, and capital markets clear and the economy faces an aggregate resource constraint of the form

$$C_t + I_t + G_t = Y_t. \quad (11)$$

Our model economy evolves along stochastic growth path. Output  $Y_t$ , consumption  $C_t$ , investment  $I_t$ , physical capital  $\bar{K}_t$  and effective capital  $K_t$  all grow at the rate  $Z_t$ . Hours worked  $L_t$  are stationary. The model can be rewritten in terms of detrended variables. We find the steady states for the detrended variables and use the method in Sims (2002) to construct a log-linear approximation of the model solution around the steady state (see Appendix). We collect all the DSGE model parameters in the vector  $\theta$ , stack the structural shocks in the vector  $\epsilon_t$ , and derive a state-space representation for the  $n \times 1$  vector  $\Delta y_t$ :

$$\Delta y_t = [\Delta \ln Y_t, \Delta \ln C_t, \Delta \ln I_t, \ln L_t]',$$

where  $\Delta$  denotes the temporal difference operator.

## 2.2 Empirical Findings

We begin by specifying a prior distribution for the parameters of the DSGE model, which is summarized in the first columns of Table 1. We are assuming that the parameters are *a priori* independent. All parameter ranges refer to 90% credible intervals. The labor share lies between 0.17 and 0.50 and the annualized growth rate of the economy ranges from 0.5 to 3.5%, which is consistent with pre-sample evidence. Our prior for the habit persistence parameter  $h$  is centered at 0.7, which is the value used by Boldrin, Christiano, and Fisher (2001). These authors find that  $h = 0.7$  enhances the ability of a standard DSGE model to account for key asset market statistics. The 90% interval for the prior distribution on  $\nu_l$  implies that the Frisch labor supply elasticity lies between 0.3 and 1.3, reflecting the micro-level estimates at the lower end, and the estimates of Kimball and Shapiro (2003) and Chang and Kim (2006) at the upper end.

The prior for the adjustment cost parameter  $s'$  is consistent with the values that Christiano, Eichenbaum, and Evans (2005) use when matching DSGE impulse response functions to consumption and investment, among other variables, to VAR responses. The prior for  $a''$  implies that in response to a 1% increase in the return to capital, utilization rates rise by 0.1 to 0.3%. These numbers are considerably smaller than the one used by Christiano, Eichenbaum, and Evans (2005). The prior on the labor adjustment cost  $\Phi$  parameter ranges from 9 to 55 and is taken from Chang, Doh, and Schorfheide (2006) who provide some justification for the numerical values. We use beta-distributions roughly centered at 0.9 to obtain a prior

for the autocorrelation parameters. Finally, the priors for the standard deviations of the structural shocks are chosen to ensure that the prior predictive distribution for the sample moments of the endogenous variables is commensurable with the magnitudes in the sample.

Figure 1 shows pointwise 90% credible bands for the predictive distribution of smoothed periodograms of the great ratios and hours worked (all series have been converted into logs). For each parameter draw from the prior (posterior) distribution, we generate a sample of 300 observations starting from the model's steady state, discard the first 100 observations, and compute a parametric spectral estimate by fitting an AR(4) model and conditioning on its least squares estimates.<sup>1</sup> Moreover, we display the (parametric) sample spectrum computed from actual U.S. data. The spectral estimates are computed after the samples have been normalized to have unit (sample) variance. The results indicate that the DSGE model is unable to explain the low frequency movements of the consumption-output ratio.

We proceed by generating draws from the posterior distribution of the DSGE model parameters using Markov Chain Monte Carlo (MCMC) techniques described in Schorfheide (2000) and An and Schorfheide (2006). Moments and 90% credible intervals for the structural parameters are provided in Table 1. While Posterior (I) is obtained from the benchmark prior distribution reported in the table, we also compute a second posterior under the restriction that the autocorrelation parameters are fixed at 0.9. With the exception of the labor adjustment parameter  $\Phi$ , the standard deviation of the labor supply shock, and the autocorrelation parameters, the two sets of posterior estimates are very similar. In the unrestricted specification, the  $\rho$ -estimates are close to unity. If the autocorrelation of the labor supply shock is restricted to be 0.9, the estimated labor adjustment cost rises to capture the persistence of hours worked. Since the adjustment costs dampen the fluctuations in hours, a more volatile labor supply shock is needed to explain the observed hours movements. In general, large autocorrelation estimates can have two interpretations. First, it could indeed be the case that preference and technology shifts are highly persistent. Second, it is possible that the exogenous shocks capture to some extent low frequency misspecifications of the DSGE model. The second column of panels in Figure 1 depicts bands for the posterior predictive distribution of sample spectra. Most strikingly, even with autocorrelation parameters near

---

<sup>1</sup>We also considered a non-parametric approach, using a Blackman-Tukey Kernel estimate with a lag window of  $M = 60$ .

unity, the DSGE model is not able to capture the persistence of the consumption-output ratio. In the next two sections we will discuss econometric techniques that allow us to relax the restrictions generated by the DSGE model. The main innovation in this paper is a method described in Section 4, which enables us to deviate from the theoretical model to different degrees at different frequencies.

### 3 Using DSGE-VARs to Compare Models and Data

We begin by defining some notation. We use the vector  $\theta$  to denote the structural parameters of the DSGE model. We assume that the DSGE model has been solved with a linear or nonlinear solution technique. While we do not take a stand on the pros and cons of linear versus nonlinear approximations, many of the procedures that we describe below are easier to implement if the structural model is solved with linear techniques.

DSGE models are tightly linked to vector autoregressions which have emerged as one of the workhorses of empirical macroeconomics in the past two decades. More specifically, DSGE models impose restrictions on vector autoregressive representations of the data. Consider the following VAR(p) model

$$y_t = \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + u_t, \quad (12)$$

where  $y_t$  is an  $n \times 1$  vector of observables and  $u_t$  is a vector of reduced-form disturbances with distribution  $u_t \sim \mathcal{N}(0, \Sigma)$ . To simplify the exposition we abstract from intercepts and trends in the VAR specification. Define  $x'_t = [y'_{t-1}, \dots, y'_{t-p}]$  and  $\Phi = [\Phi_1, \dots, \Phi_p]'$ . Suppose conditional on the DSGE parameter vector  $\theta$  one generates a sample of  $T^*$  observations  $Y^* = [y_1^*, \dots, y_{T^*}^*]'$  from the structural model. The VAR likelihood function constructed from this artificial sample, assuming that the one-step-ahead forecast errors  $u_t$  are normally distributed with mean zero and covariance matrix  $\Sigma$ , is of the form

$$p(Y^* | \Phi, \Sigma) \propto |\Sigma|^{-T^*/2} \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T^*} \text{tr}[\Sigma^{-1} (y_t^{*'} - x_t^{*'} \Phi)' (y_t^{*'} - x_t^{*'} \Phi)] \right\}. \quad (13)$$

Rather than actually simulating observations from the DSGE model it is more attractive to consider *averages* of sample moments constructed from simulated data. If the DSGE model implies a stationary law of motion for  $y_t^*$  then let us replace the sample moments



that appear in the likelihood function by population moments and add an initial improper prior  $|\Sigma|^{-(n+1)/2}$  to obtain

$$p(\Phi, \Sigma | \theta) \propto |\Sigma|^{-(T^* + n + 1)/2} \exp \left\{ -\frac{T^*}{2} \text{tr} [\Sigma^{-1} (\Gamma_{YY}(\theta) - \Phi' \Gamma_{XY}(\theta) - \Gamma_{YX}(\theta) \Phi + \Phi' \Gamma_{XX}(\theta) \Phi)] \right\}, \quad (14)$$

where

$$\Gamma_{YY}(\theta) = \mathbb{E}_\theta^D [y_t y_t'], \quad \Gamma_{YX}(\theta) = \mathbb{E}_\theta^D [y_t x_t'], \quad \Gamma_{XX}(\theta) = \mathbb{E}_\theta^D [x_t x_t'] \quad (15)$$

are the DSGE model implied covariance matrix of  $y_t^*$  and  $x_t^*$ , conditional on  $\theta$ . Now let

$$\Phi^*(\theta) = \Gamma_{XX}^{-1}(\theta) \Gamma_{XY}(\theta), \quad \Sigma^*(\theta) = \Gamma_{YY}(\theta) - \Gamma_{YX}(\theta) \Gamma_{XX}^{-1}(\theta) \Gamma_{XY}(\theta). \quad (16)$$

The matrices  $\Phi^*(\theta)$  and  $\Sigma^*(\theta)$  define a VAR approximation of the DSGE model. By construction, the first  $p$  autocovariance matrices computed from the approximation are equal to the autocovariances of the DSGE model. Since the dimension of DSGE model parameter vector  $\theta$  is typically smaller than the dimension of the VAR parameters,  $\Phi^*(\theta)$  and  $\Sigma^*(\theta)$  can be viewed as restriction functions. Deviations from the restriction functions are interpreted as misspecifications of the DSGE model.

The VAR will play two roles in our analysis. First, using the language of indirect inference, e.g., Smith (1993) and Gouriéroux, Renault, and Monfort (1993) and more recently Gallant and McCulloch (2005), the VAR serves as an approximating model for inference about the DSGE model and its parameters.  $\Phi^*(\theta)$  and  $\Sigma^*(\theta)$  define the binding function that links VAR and DSGE model parameters. Second, the estimated VAR is of interest by itself because it can be used as a device for forecasting and policy analysis and we are able to relax the DSGE model restrictions to improve its fit.<sup>2</sup>

Now suppose we interpret (14) as a prior density for the VAR coefficients  $\Phi$  and  $\Sigma$ . This prior has the property that it is centered at the VAR approximation of the DSGE model,

---

<sup>2</sup>As is well-known from the indirect inference literature, the fact that the finite-order VAR provides only an approximation to the DSGE model does not invalidate statistical inference. However, as discussed in recent work by Chari, Kehoe, and McGrattan (2004), Christiano, Eichenbaum, and Vigfusson (2006), and Fernandez-Villaverde, Rubio-Ramirez, and Sargent (2004), in the presence of approximation error one has to be careful in drawing conclusions from the estimated VAR about the validity of dynamic equilibrium models.

defined through the restriction functions  $\Phi^*(\theta)$  and  $\Sigma^*(\theta)$ :

$$\begin{aligned}\Sigma|\theta &\sim \mathcal{IW}\left(T^*\Sigma^*(\theta), T^* - k\right) \\ \Phi|\Sigma, \theta &\sim \mathcal{N}\left(\Phi^*(\theta), \Sigma \otimes [T^*\Gamma_{XX}(\theta)]^{-1}\right).\end{aligned}\tag{17}$$

Here  $\mathcal{IW}$  denotes the Inverted Wishart distribution and  $\mathcal{N}$  the normal distribution. We denote the properly normalized density of this distribution by

$$p_{\mathcal{IW}-\mathcal{N}}\left(\Phi, \Sigma \mid \Phi^*(\theta), \Sigma^*(\theta), \Gamma_{XX}(\theta), T^*\right).\tag{18}$$

The larger  $T^*$  the more concentrated the prior distribution. The use of such a prior tilts the VAR estimates toward the restrictions implied by the DSGE model.<sup>3</sup> Building on work by Ingram and Whiteman (1994), Del Negro and Schorfheide (2004) used this prior to improve forecasting and monetary policy analysis with VARs. An alternative interpretation of (14) is that the prior allows the researcher to systematically relax the DSGE model restrictions by letting  $T^*$  decrease and study how the dynamics of the VAR changes as one allows for deviations from the restrictions. Del Negro, Schorfheide, Smets, and Wouters (2006) use the setup to study the fit of the Smets-Wouters (2003) model.

More specifically, by combining the prior (17) with the likelihood function of the VAR model (12) we can obtain a joint posterior distribution for  $\theta$ ,  $\Phi$ , and  $\Sigma$ :

$$p_{\zeta}(\theta, \Phi, \Sigma|Y) \propto p(Y|\Phi, \Sigma)p_{\zeta}(\Phi, \Sigma|\theta)p(\theta),\tag{19}$$

where we define the hyperparameter  $\zeta = T^*/(T^* + T)$ . The closer  $\zeta$  is to one, the larger the number of dummy observations relative to the actual observations, or, loosely speaking, the larger the weight on the DSGE model restrictions. The estimates of the DSGE model parameters  $\theta$  can be interpreted as minimum distance estimates that are obtained by projecting the estimated VAR parameters onto the restricted subspace traced out by  $\Phi^*(\theta)$  and  $\Sigma^*(\theta)$ . To facilitate posterior simulations it is convenient to factorize the posterior as follows:

$$p_{\zeta}(\theta, \Phi, \Sigma) = p_{\zeta}(\theta|Y)p_{\zeta}(\Phi, \Sigma|Y, \theta),\tag{20}$$

where

$$p_{\zeta}(\Phi, \Sigma|Y, \theta) = p_{\mathcal{IW}-\mathcal{N}}\left(\Phi, \Sigma \mid \Phi^*(\theta), \Sigma^*(\theta), \Gamma_{XX}(\theta), T^*\right)$$

---

<sup>3</sup>Since the prior has the property of shrinking the discrepancy between VAR estimate and restriction function to zero, the procedure is often referred to as *shrinkage* estimation.

and  $p_\zeta(\theta|Y)$  is a function of

$$p_\zeta(Y|\theta) = \int p(Y|\Phi, \Sigma)p_\zeta(\Phi, \Sigma|\theta)d(\Phi, \Sigma),$$

which can be computed analytically. The marginal likelihood of the DSGE model weight  $\zeta$

$$p(Y|\zeta) = \int p_\zeta(Y|\theta)p(\theta)d\theta \tag{21}$$

can be used to assess the overall fit of the DSGE model. Loosely speaking, the marginal likelihood summarizes the discrepancy between the DSGE model implied autocovariances of  $y_t$  and the sample autocovariances. The larger this discrepancy, the smaller the value of  $\zeta$  that maximizes the marginal likelihood function.

## 4 Dummy Observations in the Frequency Domain

Our point of departure from the existing work on DSGE model priors is the observation that the prior has the potentially undesirable feature that the DSGE model restrictions are treated equally at all frequencies. However, as we pointed out in the introduction, most DSGE models are designed for business cycle analysis and we often do not expect them to capture high frequency or long-run movements in the data. As we have documented in Section 2, and other authors have pointed out as well (e.g., Whelan (2000) and Edge, Kiley and Laforte (2005)) many of the great ratios, such as consumption-to-output or the labor share are strictly speaking not stationary as implied by standard DSGE models. Models that impose invalid long-run restrictions on the data tend to be quickly rejected against specifications that allow for a more general trend structure, such as VARs. For this reason much of the early literature has either proceeded by filtering out low frequency variation from the data prior to model estimation and evaluation or, as in Watson (1993) and Diebold, Ohanian and Berkowitz (1998), conducted the empirical analysis explicitly in the frequency domain.

### 4.1 Specification of the Prior

We will generalize the prior characterized by (14) and the associated model estimation and evaluation procedures as follows. Suppose we use the dummy observations  $Y^*$  to construct

a sample periodogram:

$$F_{YY}^*(\omega) = \frac{1}{2\pi} \sum_{h=-T^*+1}^{T^*-1} \hat{\Gamma}_h^* e^{-i\omega h} = \frac{1}{2\pi} \left( \hat{\Gamma}_0^* + \sum_{h=1}^{T^*-1} (\hat{\Gamma}_h^* + \hat{\Gamma}_h^{*\prime}) \cos \omega h \right), \quad (22)$$

where  $\hat{\Gamma}_h^* = \frac{1}{T^*} \sum_{t=h+1}^{T^*} y_t^* y_{t-h}^{*\prime}$ . The likelihood function of the dummy observations has the following frequency domain approximation (see Appendix C.1 for a derivation)

$$\tilde{p}(Y^* | \Phi, \Sigma) \propto \left( \prod_{j=0}^{T^*-1} |2\pi S_V^{-1}(\omega_j, \Phi, \Sigma)| \right)^{1/2} \exp \left\{ -\frac{1}{2} \sum_{j=0}^{T^*-1} \text{tr}[S_V^{-1}(\omega_j, \Phi, \Sigma) F_{YY}^*(\omega_j)] \right\}. \quad (23)$$

Here the  $\omega_j$ 's are the fundamental frequencies  $2\pi j/T^*$ ,  $S_V^{-1}(\omega_j, \Phi, \Sigma)$  is the inverse spectral density matrix associated with the VAR

$$S_V^{-1}(\omega, \Phi, \Sigma) = 2\pi [I - M(e^{i\omega})\Phi]\Sigma^{-1}[I - \Phi'M'(e^{-i\omega})], \quad (24)$$

and  $M(z) = [Iz, \dots, Iz^p]$ . As before in the step that lead us from (13) to (14), we now replace the sample periodogram by the spectral density matrix of the DSGE model to obtain:

$$\begin{aligned} \tilde{p}(\Phi, \Sigma | \theta) &\propto \left( \prod_{j=0}^{T^*-1} |2\pi S_V(\omega_j, \Phi, \Sigma)| \right)^{-1/2} \\ &\times \exp \left\{ -\frac{1}{2} \sum_{j=0}^{T^*-1} \text{tr}[S_V^{-1}(\omega_j, \Phi, \Sigma) S_D(\omega_j, \theta)] \right\}. \end{aligned} \quad (25)$$

The advantage of the frequency domain formulation is that we are able to introduce hyperparameters that control the tightness of the prior by frequency. Let  $\lambda(\omega)$  be a weight function such that  $\frac{1}{T^*} \sum_{j=0}^{T^*-1} \lambda(\omega_j) = 1$  (or  $\int_0^{2\pi} \lambda(\omega) d\omega = 2\pi$ ). We can modify the prior as follows:

$$\begin{aligned} \tilde{p}(\Phi, \Sigma | \theta) &\propto \exp \left\{ \frac{1}{2} \frac{T^*}{2\pi} \frac{2\pi}{T^*} \sum_{j=0}^{T^*-1} \lambda(\omega_j) \ln \left| \frac{1}{2\pi} S_V^{-1}(\omega_j, \Phi, \Sigma) \right| \right\} \\ &\times \exp \left\{ -\frac{1}{2} \frac{T^*}{2\pi} \frac{2\pi}{T^*} \sum_{j=0}^{T^*-1} \lambda(\omega_j) \text{tr}[S_V^{-1}(\omega_j, \Phi, \Sigma) S_D(\omega_j, \theta)] \right\}. \end{aligned} \quad (26)$$

Using the definition of  $S_V^{-1}(\omega, \Phi, \Sigma)$  from (24) we can rewrite the trace in (26) as follows:

$$\begin{aligned}
& \text{tr}[S_V^{-1}(\omega_j, \Phi, \Sigma)S_D(\omega_j, \theta)] \\
&= 2\pi \text{tr} \left[ \Sigma^{-1} (I - \Phi' M'(e^{-i\omega})) S_D(\omega_j, \theta) (I - M(e^{i\omega}) \Phi) \right] \\
&= 2\pi \text{tr} \left[ \Sigma^{-1} \left( S_D(\omega_j, \theta) - \Phi' M'(e^{-i\omega}) S_D(\omega_j, \theta) - S_D(\omega_j, \theta) M(e^{i\omega}) \Phi \right. \right. \\
&\quad \left. \left. + \Phi' M'(e^{-i\omega}) S_D(\omega_j, \theta) M(e^{i\omega}) \Phi \right) \right] \\
&= 2\pi \text{tr} \left[ \Sigma^{-1} \left( S_D(\omega_j, \theta) - \Phi' \text{re}(M'(e^{-i\omega})) S_D(\omega_j, \theta) - S_D(\omega_j, \theta) \text{re}(M(e^{i\omega})) \Phi \right. \right. \\
&\quad \left. \left. + \Phi' M'(e^{-i\omega}) S_D(\omega_j, \theta) M(e^{i\omega}) \Phi \right) \right].
\end{aligned}$$

Here  $\text{re}(C)$  denotes the real part of the complex matrix  $C$ . If we now replace the summations over the fundamental frequencies  $\omega_j$  in (26) by integrals, and add an initial improper prior  $\mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}} |\Sigma|^{-(n+1)/2}$ , we can obtain the following representation

$$\begin{aligned}
p(\Phi, \Sigma | \theta) &\propto \mathcal{I}_{\{\Phi \in \mathcal{P}\}} |\Sigma|^{-(T^* + n + 1)/2} f_{\lambda, T^*}(\Phi) \\
&\quad \times \exp \left\{ -\frac{T^*}{2} \text{tr} \left[ \Sigma^{-1} (\Gamma_{\lambda, Y Y}(\theta) - 2\Gamma_{\lambda, Y X}(\theta) \Phi + \Phi' \Gamma_{\lambda, X X}(\theta) \Phi) \right] \right\}, \quad (27)
\end{aligned}$$

where  $\mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}}$  is the indicator function that is one if  $\Phi \in \text{int}(\mathcal{P})$ ,  $\mathcal{P}$  is the set of parameter values for which the VAR is non-explosive, and  $\text{int}(\mathcal{P})$  denotes its interior.<sup>4</sup> Moreover,

$$f_{\lambda, T^*}(\Phi) = \exp \left\{ \frac{T^*}{2 \cdot 2\pi} \int_0^{2\pi} \lambda(\omega) \ln |(I - M(e^{i\omega}) \Phi)(I - \Phi' M'(e^{-i\omega}))| d\omega \right\},$$

and

$$\begin{aligned}
\Gamma_{\lambda, Y Y}(\theta) &= \int_0^{2\pi} \lambda(\omega) S_D(\omega, \theta) d\omega, & \Gamma_{\lambda, Y X}(\theta) &= \int_0^{2\pi} \lambda(\omega) S_D(\omega, \theta) \text{re}(M(e^{i\omega})) d\omega \\
\Gamma_{\lambda, X X}(\theta) &= \int_0^{2\pi} \lambda(\omega) M'(e^{-i\omega}) S_D(\omega, \theta) M(e^{i\omega}) d\omega.
\end{aligned} \quad (28)$$

Finally, define

$$\Phi_{\lambda}^*(\theta) = \Gamma_{\lambda, X X}^{-1}(\theta) \Gamma_{\lambda, X Y}(\theta), \quad \Sigma_{\lambda}^*(\theta) = \Gamma_{\lambda, Y Y}(\theta) - \Gamma_{\lambda, Y X}(\theta) \Gamma_{\lambda, X X}^{-1}(\theta) \Gamma_{\lambda, X Y}(\theta). \quad (29)$$

and rewrite the prior density as

$$\begin{aligned}
p(\Phi, \Sigma | \theta) &= c(\lambda, T^*, \theta) \mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}} f_{\lambda, T^*}(\Phi) \\
&\quad \times p_{\mathcal{I}W-\mathcal{N}} \left( \Phi, \Sigma \mid \Phi_{\lambda}^*(\theta), \Sigma_{\lambda}^*(\theta), \Gamma_{\lambda, X X}(\theta), T^* \right),
\end{aligned} \quad (30)$$

<sup>4</sup>Depending on the choice of  $\lambda(\omega)$ , the set  $\mathcal{P}$  can be enlarged. If  $\Lambda_{ll}$ ,  $l = 1, \dots, np$  are the possibly complex eigenvalues of  $\Phi$  (written in companion form), it has to be guaranteed that  $0 < 1 + |\Lambda_{ll}|^2 - 2\text{re}(\Lambda_{ll}) \cos(\omega) - 2\text{im}(\Lambda_{ll}) \sin(\omega)$  for all  $\omega$  with  $\lambda(\omega) > 0$ .

where  $p_{\mathcal{I}\mathcal{W}-\mathcal{N}}(\cdot)$  was defined in (18) and  $c(\lambda, T^*, \theta)$  ensures that the density function is properly normalized.

**Remark:** In the special case of  $\lambda(\omega) = 1$  the matrices  $\Gamma_{\lambda, \cdot}(\theta)$  reduce to the time domain counterpart given in (15). Moreover, (see Appendix B.2) since

$$\int_0^{2\pi} \ln \left| \frac{1}{2\pi} S_V^{-1}(\omega, \Phi, \Sigma) \right| d\omega = -2\pi \ln |\Sigma|$$

it follows that  $f_{\lambda, T^*}(\Phi) = 1$  for all  $\Phi$  and  $T^*$ . Hence, the prior density in (27) reduces to its time domain analogue (14) and the prior takes the familiar  $\mathcal{I}\mathcal{W} - \mathcal{N}$  form.  $\square$

As in the previous section, we introduce the hyperparameter  $0 \leq \zeta \leq 1$  to control the overall degree of shrinkage:  $\zeta = T^*/(T^* + T)$ , where  $T$  is the size of the actual sample that is used to estimate the model. The prior  $p(\Phi, \Sigma|\theta)$  can now be combined with a prior distribution for the DSGE model parameters,  $p(\theta)$ , and the VAR-based likelihood function constructed from the a sample of actual observations  $Y$ , denoted by  $\mathcal{L}(\Phi, \Sigma|Y)$ , to conduct Bayesian inference.

Our proposed procedure differs from a Bayesian version of band-spectrum regression in that all the frequencies are used (and equally weighted) in the construction of the likelihood function. Hence, the estimated DSGE-VAR can be used to forecast short-run fluctuations as well as long-run trends. The key feature of our analysis is that the degree of shrinkage toward the DSGE model restrictions, determined by  $\lambda(\omega)$ , can be frequency-specific. Suppose that  $\lambda(\omega)$  is large a business cycle frequencies and zero elsewhere. The resulting prior will penalize VAR estimates that imply large discrepancies between the spectrum of the DSGE model and the spectrum of the VAR at business cycle frequencies.

## 4.2 Posterior Distributions

We begin by characterizing the posterior distribution conditional on the DSGE model parameters  $\theta$ . The likelihood function is of the form

$$p(Y|\Phi, \Sigma) = (2\pi)^{-nT/2} |\Sigma|^{-T/2} \exp \left\{ -\frac{T}{2} \text{tr} [\Sigma^{-1} (\hat{\Gamma}_{YY} - 2\hat{\Gamma}_{YX}\Phi + \Phi' \hat{\Gamma}_{XX} \Phi)] \right\}, \quad (31)$$

where, for instance,  $\hat{\Gamma}_{YY}$  denotes the sample moment  $\frac{1}{T} \sum y_t y_t'$ . We deduce from Bayes Theorem

$$p(\Phi, \Sigma | Y, \theta) \propto c(\lambda, T^*, \theta) \mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}} f_{\lambda, T^*}(\Phi) \times \exp \left\{ -\frac{T^* + T}{2} \text{tr} \left[ \Sigma^{-1} \left( \tilde{\Gamma}_{\lambda, \zeta, YY}(\theta) - 2\tilde{\Gamma}_{\lambda, \zeta, YX}(\theta)\Phi + \Phi' \tilde{\Gamma}_{\lambda, \zeta, XX}(\theta)\Phi \right) \right] \right\}, \quad (32)$$

using the notation that  $\tilde{\Gamma}_{\lambda, \zeta, YY}(\theta) = \zeta \Gamma_{\lambda, YY}(\theta) + (1 - \zeta) \hat{\Gamma}_{YY}$ . As before, we define

$$\begin{aligned} \tilde{\Phi}_{\lambda, \zeta}(\theta) &= \tilde{\Gamma}_{\lambda, \zeta, XX}^{-1}(\theta) \tilde{\Gamma}_{\lambda, \zeta, XY}(\theta), \\ \tilde{\Sigma}_{\lambda, \zeta}(\theta) &= \tilde{\Gamma}_{\lambda, \zeta, YY}(\theta) - \tilde{\Gamma}_{\lambda, \zeta, YX}(\theta) \tilde{\Gamma}_{\lambda, \zeta, XX}^{-1}(\theta) \tilde{\Gamma}_{\lambda, \zeta, XY}(\theta). \end{aligned}$$

and can write the posterior density as

$$p(\Phi, \Sigma | Y, \theta) = c(\lambda, T^*, \theta) \mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}} f_{\lambda, T^*}(\Phi) \times p_{\mathcal{TW-N}} \left( \Phi, \Sigma \left| \tilde{\Phi}_{\lambda, \zeta}(\theta), \tilde{\Sigma}_{\lambda, \zeta}(\theta), \tilde{\Gamma}_{\lambda, \zeta, XX}(\theta), T^* + T \right. \right). \quad (33)$$

**Remark:** If  $\lambda(\omega) = 1$  then the adjustment term  $f_{\lambda, T^*}(\Phi) = 1$  and we can use Algorithm 1 to generate parameter draws from the posterior. In the general case of  $\lambda(\omega) \neq 1$  the posterior distribution of  $\Phi$  conditional on  $\Sigma$  and  $\theta$  is non-standard and the normalizing constant of the prior density cannot be calculated analytically.  $\square$

### 4.3 Discussion

**Bandpass-filtered Dummy Observations.** Suppose we use bandpass-filtered dummy observations to construct a prior distribution instead of the approach outlined in the previous section. Assume that the bandpass filter has a transfer function of the form

$$B(e^{-i\omega})B'(e^{i\omega}) = |B(e^{-i\omega})|^2 = I \cdot \lambda(\omega), \quad (34)$$

where  $B(\cdot)$  is a diagonal matrix. Let  $S_D(\omega, \theta)$  be the spectrum of the DSGE model generated observations and define

$$S_D^B(\omega, \theta) = B(e^{-i\omega})S_D(\omega, \theta)B'(e^{i\omega}) = \lambda(\omega)S_D(\omega, \theta) \quad (35)$$

as the spectrum of the filtered observations. Then the prior constructed from the filtered dummy observations can be represented as

$$p(\Phi, \Sigma | \theta) \propto \mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}} \times p_{\mathcal{TW-N}} \left( \Phi, \Sigma \left| \Phi_{\lambda}^*(\theta), \Sigma_{\lambda}^*(\theta), \Gamma_{\lambda, XX}(\theta), T^* \right. \right), \quad (36)$$

which is identical to (27) with the exception that the adjustment term  $f_{\lambda, T^*}(\Phi)$  is absent.

**Relationship to Band Spectrum Regression.** The restriction function  $\Phi_\lambda^*(\theta)$  can be viewed as the population analog of a band spectrum regression estimator of  $\Phi$  (see Engle (1974)), constructed from the dummy observations. Let  $Y^*$  and  $X^*$  be composed of (un-filtered) dummy observations from the DSGE model. Let  $W$  be the  $T^* \times T^*$  matrix with elements

$$W_{j,t} = \frac{1}{\sqrt{T^*}} e^{i\omega_j t}$$

We use  $\dagger$  to denote the complex conjugate of the transpose of a matrix. Moreover,  $\Lambda$  is a  $T^* \times T^*$  diagonal matrix with entries  $\lambda^{1/2}(\omega_j)$ , which re-weights different frequencies. Then the band-spectrum estimator of  $\Phi$  in the VAR  $Y^* = X^*\Phi + U$  is given by

$$\begin{aligned} \widehat{\Phi}_B &= (X^{*\prime} W^\dagger \Lambda' \Lambda W X^*)^{-1} X^{*\prime} W^\dagger \Lambda' \Lambda W Y^* \\ &= \left( \frac{1}{T^*} \sum_{j=0}^{T^*-1} \lambda(\omega_j) F_{XX}^*(\omega_j) \right)^{-1} \frac{1}{T^*} \sum_{j=0}^{T^*-1} \lambda(\omega_j) F_{XY}^*(\omega_j). \end{aligned}$$

and converges to  $\Phi_\lambda^*(\theta)$  [needs to be verified]. Here  $F_{XX}^*(\omega_j) = (WX)_{\cdot j}^\dagger (WX)_j$  and  $F_{XY}^*(\omega_j) = (WX)_{\cdot j}^\dagger (WY)_j$  denote sample cross periodograms. Hence, the prior constructed from bandpass-filtered dummy observations is centered at the (population) band-spectrum regression estimator of  $\Phi$ . As shown in Engle (1980), this estimator is in general not a consistent estimator of the value of  $\Phi$  that locally approximates the target spectral density  $S_D(\omega, \theta)$  if frequency bands are omitted by setting certain  $\lambda(\omega_j)$ 's equal to zero.

Alternatively, consider the mode of the prior developed in Section 4. Let  $\psi = [\text{vec}(\Phi)^\dagger, \text{vech}(\Sigma)^\dagger]^\dagger$  and denote the mode of the prior by  $\tilde{\psi}$ . At the mode, the following first-order conditions are satisfied (for all  $j$ )

$$0 = \int \lambda(\omega) \text{tr} \left[ \left( S_V(\omega, \tilde{\Phi}, \tilde{\Sigma}) - S_D(\omega, \theta) \right) \frac{\partial S_V^{-1}(\omega, \tilde{\Phi}, \tilde{\Sigma})}{\partial \psi_j} \right] d\omega = 0.$$

Hence, at the prior mode we minimize a weighted discrepancy between the spectral density of the DSGE model and the VAR. Notice that in general the prior does not peak at the band-spectrum estimate, the exception being the case in which at the band-spectrum estimate [check this]

$$S_V(\omega, \widehat{\Phi}_B, \widehat{\Sigma}_B) = S_D(\omega, \theta) \quad \text{whenever} \quad \lambda(\omega) > 0.$$



**Intercepts, Trends, and Nonstationarities.** The VAR(p) model in (12) was specified without intercept and trend component, which are important in applications. To include deterministic trends we re-write the VAR as follows:

$$y_t = \Psi_0 + \Psi_1 t + \tilde{y}_t, \quad \tilde{y}_t = \Phi_1 \tilde{y}_{t-1} + \dots + \Phi_p \tilde{y}_{t-p} + u_t. \quad (37)$$

The specification of (37) is consistent with the DSGE model. The intercept  $\Psi_0$  captures model implied steady-state ratios for the observables, and the trend term  $\Psi_1 t$  picks up deterministic trend components, induced, for instance, by the drift in the random walk technology process of the model outlined in Section 2 or simply by a deterministic labor augmenting trend. In our subsequent application, we will apply the dummy observation prior to the autoregressive coefficient matrices  $\Phi_1, \dots, \Phi_p$ , and use a separate prior, also centered at the DSGE model predictions, for the coefficient matrices  $\Psi_0$  and  $\Psi_1$ .

So far we assumed that the DSGE model implies that  $y_t^*$ , or  $\tilde{y}_t^*$  in the notation of (37), is stationary. However, many macroeconomic time series including output, consumption, and investment, are highly persistent and often better characterized as difference stationary processes. Non-stationary behavior of endogenous variables in DSGE models is typically generated by assuming that some of the exogenous processes, for instance the technology process, have unit roots. If some elements of  $y_t^*$  are difference-stationary then the autocovariance matrices that appear in (15) are not defined. Del Negro, Schorfheide, Smets, and Wouters (2006) circumvent the problem by rewriting the VAR in vector error correction (VECM) form. However, the VECM specification has a major disadvantage: it dogmatically imposes the DSGE model's potentially misspecified common trend restrictions onto the VAR representation.

The frequency domain dummy observation approach allows for much more flexibility. Suppose we start from the spectrum for  $\Delta y_t$ , denoted by  $S_D^\Delta(\omega)$ . Let  $D(z) = I(1 - z)$  be the difference filter such that its inverse “integrates”  $\Delta y_t$ . Then we can define

$$S_D(\omega, \theta) = D^{-1}(e^{-i\omega}) S_D^\Delta(\omega, \theta) D^{-1'}(e^{i\omega}) = \frac{1}{2 - 2 \cos \omega} S_D^\Delta(\omega, \theta).$$

As long as  $\lambda(\omega)$  is zero in a neighborhood of  $\omega = 0$ , the quasi-spectral density  $S_D(\omega, \theta)$  and hence the restriction functions  $\Phi^*(\theta)$  and  $\Sigma^*(\theta)$  are well defined for a vector autoregressive model that is specified in terms of the levels of  $y_t$ . By putting little weight on near zero

frequencies we can assign less weight on the common trend restrictions of the DSGE model to account for non-stationarities of the great ratios in the data and more weight on its business cycle implications.

**A Modified Prior Distribution.** From a computational perspective the proposed prior density is rather awkward. The normalization constant is unknown and it is not possible to generate independent draws from the prior. As an alternative, we will consider a prior for  $\Phi$  that is Gaussian conditional on  $\Sigma$ , based on a quadratic approximation of the log adjustment term  $\ln f_{\lambda, T^*}(\Phi)$ . This approximation is provided in Appendices B.3 and B.4.

## 5 Examples

This section provides two numerical examples that illustrate some of the features of the proposed prior distribution. The first example consists of a prior distribution for an AR(1) model, that is derived from a target spectral density that corresponds to the sum of two AR(1) process with different degrees of autocorrelation. We consider three weight functions  $\lambda(\omega)$ , generate parameter draws from the prior distribution, and show how the implied spectral density changes as a function of  $\lambda(\omega)$ . In the second example we consider a bivariate vector autoregression. We estimate the VAR under the frequency domain dummy observation prior and compare the implied posterior distribution of the spectrum under various weight functions  $\lambda(\omega)$ . The data used in the estimation of the VAR are generated from a process that relative to the target spectral density  $S_D(\omega)$  has an additional low frequency component, which renders  $S_D(\omega)$  misspecified at low frequencies. We also compute marginal data densities for the VAR under the various prior distributions.

### 5.1 An AR(1) Model

Consider the simple AR(1) model  $y_t = \phi y_{t-1} + u_t$  with spectral density function

$$S_V(\omega, \phi, \sigma) = \frac{1}{2\pi} \frac{\sigma^2}{1 + \phi^2 - 2\phi \cos \omega}. \quad (38)$$

We assume that the DSGE model does not depend on any unknown parameters and hence let  $S_D(\omega, \theta) = S_D(\omega)$ . From (27) it is straightforward to verify that the mode of the prior

distribution,  $[\tilde{\phi}, \tilde{\sigma}]'$ , minimizes the weighted discrepancy between the AR(1) implied spectral density and the DSGE model spectral density function, that is,

$$[\tilde{\phi}, \tilde{\sigma}]' = \operatorname{argmin}_{\phi, \sigma} \int \frac{\lambda(\omega)}{S_V^2(\omega, \tilde{\phi}, \tilde{\sigma})} [S_V(\omega, \phi, \sigma) - S_D(\omega)]^2 d\omega.$$

Thus, the prior density implicitly penalizes parameterizations of the AR(1) model that yield spectral densities that are very different from that implied by the DSGE model.

Now define the weighted spectrum of  $y_t$  and the cross-spectrum of  $y_t$  and  $y_{t-1}$

$$\gamma_{\lambda,0} = \int_0^{2\pi} \lambda(\omega) S_D(\omega) d\omega, \quad \gamma_{\lambda,1} = \int_0^{2\pi} \lambda(\omega) \cos(\omega) S_D(\omega) d\omega.$$

The prior distribution (27) therefore simplifies to

$$p(\phi, \sigma^2) = c(\lambda, T^*) \mathcal{I}_{\{|\phi| < 1\}} f_{\lambda, T^*}(\phi) p_{\mathcal{IG}-\mathcal{N}}\left(\phi, \sigma^2 \mid \phi_\lambda^*, \sigma_\lambda^{*2}, \gamma_{\lambda,0}, T^*\right), \quad (39)$$

where

$$\phi_\lambda^* = \gamma_{\lambda,0}^{-1} \gamma_{\lambda,1}, \quad \sigma_\lambda^{*2} = \gamma_{\lambda,0} - \gamma_{\lambda,1}^2 / \gamma_{\lambda,0}.$$

and

$$f_{\lambda, T^*}(\phi) = \exp \left\{ \frac{T^*}{2 \cdot 2\pi} \int_0^{2\pi} \lambda(\omega) \ln(1 + \phi^2 - 2\phi \cos \omega) d\omega \right\}.$$

We can generate dependent draws from the prior distribution using a Metropolis-within-Gibbs algorithm.

**Algorithm 1:** *MCMC Algorithm for Prior Distribution.* For  $s = 1$  to  $n_{sim}$  iterate over the following two steps:

1. Draw  $\sigma^{(s)}$  conditional on  $\phi^{(s-1)}$  from an inverse Gamma distribution:

$$\sigma^{(s)} \sim \mathcal{IG}\left(T^* (1 + \phi^{(s-1)})^2 \gamma_{\lambda,0} - 2\phi^{s-1} \gamma_{\lambda,1}, T^*\right).$$

2. Draw  $\vartheta$  from a normal distribution  $\mathcal{N}(\phi^{(s-1)}, \sigma_{(s)}^2 [T^* \gamma_{\lambda,0}]^{-1})$ . Let

$$\phi^{(s)} = \begin{cases} \vartheta & \text{with probability } \min \left[ 1, \frac{p(\vartheta, \sigma_{(s)}^2)}{p(\phi^{(s-1)}, \sigma_{(s)}^2)} \right] \\ \phi^{(s-1)} & \text{otherwise} \end{cases}.$$

Here  $p(\phi, \sigma)$  is given in (39).  $\square$

To illustrate the properties of this prior distribution we provide a numerical example.

Let

$$S_D(\omega) = \frac{1}{2\pi} \frac{1}{1 + 0.5^2 - 2 \cdot 0.5 \cos(\omega)} + \frac{1}{2\pi} \frac{0.05}{1 + 0.9^2 - 2 \cdot 0.9 \cos(\omega)}. \quad (40)$$

Hence,  $S_D(\omega)$  is the spectral density matrix associated with the sum of two AR(1) processes with different degrees of autocorrelation.

Parameter draws are plotted in Figure 2, whereas Figure 3 depicts 90% bands for draws of the implied spectral density functions. The (1,1) panels correspond to the benchmark case of  $\lambda(\omega) = 1$ . The weight function for (1,2) emphasizes the low frequencies whereas the  $\lambda(\omega)$ 's in panels (2,1) and (2,2) amplify the high frequencies. While the prior means of the parameters are fairly similar in all four cases, the correlation between  $\phi$  and  $\sigma$  differs substantially. There is a strong negative correlation if the low frequencies are heavily weighted, whereas the correlation is slightly positive if emphasis is placed on the high frequencies. We see in panels (2,1) and (2,2) that the prior places a lot of weight on spectral densities that match the target spectrum  $S_D(\omega)$  at high frequencies. At the same time, the low frequency behavior is allowed to deviate substantially from the target density. The picture reverses if we use a weight function that emphasizes low frequencies, as can be seen from Panel (1,2) of Figure 3.

The drawback of our prior is that due to the adjustment term the normalization constant cannot be calculated analytically. Knowledge of the normalization constant is important to compute marginal data densities and use the prior in a hierarchical setting in which the target spectral density matrix is indexed by a parameter  $\theta$ . We consider an alternative prior, which we refer to as “approximate,” in which we approximate the conditional density

$$p(\phi|\sigma^2) \propto \mathcal{I}_{\{|\phi|<1\}} f_{\lambda, T^*}(\phi) p_{\text{IG-N}}\left(\phi, \sigma^2 \mid \phi_{\lambda}^*, \sigma_{\lambda}^{*2}, \gamma_{\lambda, 0}, T^*\right),$$

by a normal density. More specifically, we approximate  $\ln p(\phi|\sigma^2)$  by a quadratic function of  $\phi$  around the mode  $\tilde{\phi}(\sigma^2) = \operatorname{argmax}_{\phi} \ln p(\phi|\sigma^2)$ . Details of this approximation are provide in Appendix C.3. Parameter and spectral density draws from the prior distribution are plotted in Figures 4 and 5. These draws look very similar to the ones obtained under the “exact” prior and have the same qualitative features.

Finally, we plot draws from the prior (of the parameters and the spectral densities) obtained if we use the bandpass-filtered dummy observations, ignoring the term  $f_{\lambda, T^*}(\phi)$  in

Figures 6 and 7. It turns out that the prior is quite different from the one that is obtained if the adjustment term from the frequency domain likelihood function is included due to the inconsistency of the band spectrum estimator in dynamic models as discussed in Engle (1980). In particular, the implied prior of the spectrum from the bandpass-filtered dummy observations does not always concentrate near the target spectrum in areas of the spectral bands where the weight function  $\lambda(\omega)$  is large. Engle (1980, p. 400) provides some analytical calculations for the AR(1) model.

## 5.2 A Bivariate VAR

Let  $y_t$  now be a  $2 \times 1$  vector such as consumption and investment. Suppose that according to a DSGE model the short-run dynamics of  $y_t$  are described by the following detrended variables

$$\tilde{y}_t = \Psi \tilde{y}_{t-1} + u_t. \quad (41)$$

Hence, the spectrum is given by

$$S_D(\omega) = \frac{1}{2\pi} (I - \Psi e^{-i\omega})^{-1} \Sigma_u (I - \Psi' e^{i\omega})^{-1}. \quad (42)$$

Suppose according to the DGP there is a stochastic trends that influences consumption and investment:

$$x_t = \rho x_{t-1} + \eta_t. \quad (43)$$

According to the DGP the relationship between the observables  $y_t$ , the detrended variables  $\tilde{y}_t$  and the trends  $x_t$  is of the form

$$y_t = \Xi x_t + \tilde{y}_t, \quad (44)$$

where  $\Xi = [1, 1]'$ . Moreover, we assume that  $\eta_t$  and  $u_t$  are independent at all leads and lags.

The “true” spectrum of  $y_t$  is therefore given by

$$S_y(\omega) = \frac{1}{2\pi} (I - \Psi e^{-i\omega})^{-1} \Sigma_u (I - \Psi' e^{i\omega})^{-1} + \frac{1}{2\pi} \frac{\sigma_\eta^2}{1 + \rho^2 - 2\rho \cos(\omega)} \Xi \Xi'. \quad (45)$$

We consider the following parameterization:

$$\Psi = \begin{bmatrix} 0.7 & 0.3 \\ -0.1 & 0.8 \end{bmatrix}, \quad \Sigma_u = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix}, \quad \rho = 0.98, \quad \sigma_\eta^2 = 0.1.$$

Figure 8 depicts the spectral densities for the (misspecified) DSGE model and the DGP. Under the DGP and the DSGE the spectra peak at the origin due to the near random walk component. The spectrum of the detrended DSGE matches that of the DGP and the non-detrended DGP for frequencies  $\omega > 0.08\pi$ .

We proceed by specifying the weight function  $\lambda(\omega)$ . Frequencies below  $\omega = 0.001$  are suppressed:  $\lambda(\omega) = 0$ . The low frequency band  $\omega \in [0.001, 0.08\pi]$  are scaled by  $\lambda_1$ , and all other frequencies (business cycle and higher frequencies) are scaled by  $\lambda_2$ . Since the weights have to normalize to one, we parameterize the step function in terms of  $\lambda = \lambda_1/\lambda_2$  and consider three values  $\lambda \in \{1/10, 1, 10\}$ . Moreover, we set  $T^* = 120$ .

Using the frequency domain dummy observations we now construct a prior distribution for a bivariate VAR with  $p = 4$  lags. We are generating draws from this prior distribution using a Metropolis-within-Gibbs algorithm for four different choices of  $\lambda(\omega)$ .

**Algorithm 2:** *MCMC Algorithm for Prior Distribution.* For  $s = 1$  to  $n_{sim}$  iterate over the following two steps:

1. Draw  $\Sigma^{(s)}$  conditional on  $\Phi^{(s-1)}$  from an inverse Wishart distribution:

$$\Sigma^{(s)} \sim \mathcal{IW}\left(T^*(\Gamma_{\lambda,YY} - 2\Gamma_{\lambda,YX}\Phi + \Phi'\Gamma_{\lambda,XX}\Phi), T^*\right).$$

2. Draw  $\vartheta$  from a normal distribution  $\mathcal{N}(\Phi^{(s-1)}, \Sigma^{(s)} \otimes [T^*\Gamma_{\lambda,XX}]^{-1})$ . Let

$$\phi^{(s)} = \begin{cases} \vartheta & \text{with probability } \min\left[1, \frac{p(\vartheta, \Sigma^{(s)})}{p(\Phi^{(s-1)}, \Sigma^{(s)})}\right] \\ \Phi^{(s-1)} & \text{otherwise} \end{cases}.$$

Here  $p(\Phi, \Sigma)$  is given in (27).  $\square$

Parameter draws from the prior distribution are converted into spectral densities and are plotted in Figure 9. We also depict the weight functions  $\lambda(\omega)$  and the spectral densities of the DSGE model  $S_D(\omega)$  for the two elements of  $y_t$ . As in Example 1, the prior is fairly diffuse on the low frequency behavior for  $\lambda = 1/10$ . Vice versa, if we set  $\lambda = 10$ , the spectral density draws are tightly concentrated around  $S_D(\omega)$  for  $\omega < 0.08\pi$ .

We now simulate  $T = 120$  observations from the data generating process (44) and generate draws from the posterior distribution of the VAR(4) using a modified version of Algorithm 2. This implies that the hyperparameter  $\zeta = 0.5$ .

**Algorithm 3:** *MCMC Algorithm for Posterior Distribution.* Obtained by straightforward modification of Algorithm 2 based on Equations (32).

Figure 10 depicts draws from the posterior distribution of the spectral densities. For  $\lambda = 1/10$  (top panel) our prior shrinks only toward the correctly specified business cycle / high frequency restrictions of the DSGE model. Hence, in the posterior distribution we are able to correctly pick up the low frequency behavior of the DGP. As the weight on the low frequency restrictions is increased (middle and bottom panels), the VAR estimates more and more reflect the misspecified low frequency behavior of the DSGE model. Marginal data densities are reported in Table 2.

## 6 DSGE Model Application

So far:

- Condition on the posterior mean estimate of  $\theta$  obtained in Section 2, under the prior that fixes  $\rho_g = \rho_\phi = 0.9$  (Posterior (II) in Table 1). The joint estimation of the DSGE model and VAR parameters is not yet operational.
- The VAR has 4 lags and is specified in log levels of output, consumption, investment, and hours. All variables are scaled by 100, such that log differences can be interpreted as quarter-to-quarter percentage changes.
- We use Algorithm 2 to generate draws from the prior distribution of the VAR parameters. For each draw, we simulate 300 observations from the estimated VAR, using actual U.S. data from QIV:2005 to initialize the VAR lags for the estimation. The first 100 draws are discarded, and we construct univariate parametric spectral density estimates for the simulated data based on estimated AR(4) models. Before computing the density estimates, we standardize the simulated samples to have variance one. We consider the following series: output growth, consumption growth, investment growth, log hours worked, log consumption-output ratio, and log investment-output ratio.
- Figures 11, 12, and 13 depicts the DSGE-VAR prior implied distribution of the sample spectral densities together with the actual sample densities. We use the class of

weight functions  $\lambda(\omega)$  described in Section 5.2. The prior that weighs all frequencies equally looks similar to the one that emphasizes the long-run frequencies. For output, consumption, and investment growth the prior works as expected: if we emphasize the business cycle frequencies then the prior predictive distribution becomes more diffuse at the low frequencies. Unfortunately, this effect is less pronounced for the great ratios and hours worked.

## 7 Conclusions

(to be written)

## References

- Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Mattias Villani (2006): “Evaluating an Estimated New Keynesian Small Open Economy Model,” *Manuscript*, Sveriges Riksbank.
- Calvo, Guillermo (1983): “Staggered Prices in a Utility-Maximizing Framework,” *Journal of Monetary Economics*, **12**, 383-398.
- Chari, V.V., Patrick Kehoe, and Ellen McGrattan (2004): “A Critique of Structural VARs Using Business Cycle Theory,” *Manuscript*, Federal Reserve Bank of Minneapolis.
- Christiano, Lawrence, Martin Eichenbaum, and Charles Evans (2005): “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, **113**, 1-45.
- Christiano, Lawrence, Martin Eichenbaum, and Robert Vigfusson (2006): “Assessing Structural VARs,” *NBER Macroeconomics Annual*, forthcoming.
- Del Negro, Marco and Frank Schorfheide (2004): “Priors from General Equilibrium Models for VARs,” *International Economic Review*, **45**, 643-673.
- Del Negro, Marco and Frank Schorfheide (2005): “Monetary Policy Analysis with Potentially Misspecified Models,” *Manuscript*, University of Pennsylvania.



- Del Negro, Marco, Frank Schorfheide, Frank Smets, and Raf Wouters (2006): “On the Fit and Forecasting Performance of New Keynesian Models,” *Manuscript*, University of Pennsylvania.
- Diebold, Francis, Lee Ohanian, and Jeremy Berkowitz (1998): “Dynamic Equilibrium Economies: A Framework for Comparing Models and Data,” *Review of Economic Studies*, **65**, 433-452.
- Edge, Rochelle, Michael Kiley, and Jean-Philippe Laforge (2005): “An Estimated DSGE Model of the US Economy,” *Manuscript*, Board of Governors.
- Engle, Robert F. (1974): “Band Spectrum Regression,” *International Economic Review*, **15**, 1-11.
- Engle, Robert F. (1980): “Exact Maximum Likelihood Methods for Dynamic Regressions and Band Spectrum Regressions,” *International Economic Review*, **21**, 391-407.
- Espasa, A. (1977): “*The Spectral Maximum Likelihood Estimation of Econometric Models with Stationary Errors*,” Vandenhoeck and Ruprecht, Göttingen.
- Fernández-Villaverde, Jesús, Juan Rubio-Ramírez, and Thomas Sargent (2004): “A, B, C’s (and D’s) for Understanding VARs,” *Manuscript*, University of Pennsylvania.
- Gallant, A. Ronald and Robert E. McCulloch (2005): “On the Determination of General Scientific Models,” *Manuscript*, Duke University.
- Geweke, John (1999): “Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication,” *Econometric Reviews*, **18**, 1-126.
- Gourieroux, Christian, Eric Renault, and Alain Monfort (1993): “Indirect Inference,” *Journal of Applied Econometrics*, **8**, S85-S118.
- Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1998): “Long-Run Implications of Investment-Specific Technological Change,” *American Economic Review*, **87**(3), 342-362.
- Ingram, Beth and Charles Whiteman (1994): “Supplanting the Minnesota prior – Forecasting macroeconomic time series using real business cycle model priors,” *Journal of Monetary Economics*, **34**, 497-510.

- King, Robert G., Charles I. Plosser, and Sergio T. Rebelo (1988): "Production, Growth, and Business Cycles II: New Directions," *Journal of Monetary Economics*, **21**, 309-341.
- Klein, Lawrence R. and Richard F. Kosobud (1961): "Some Econometrics of Growth: Great Ratios of Economics," *Quarterly Journal of Economics*, **75**(2), 173-198.
- Schorfheide, Frank (2000): "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, **15**, 645-670.
- Sims, Christopher (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, **20**, 1-20.
- Smets, Frank and Raf Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model for the Euro Area," *Journal of the European Economic Association*, **1**, 1123-1175.
- Smith, Anthony (1993): "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, **8**, S63-S84.
- Theil, Henry and Arthur S. Goldberger (1961): "On Pure and Mixed Estimation in Economics". *International Economic Review*, **2**, 65-78.
- Watson, Mark (1993): "Measures of Fit for Calibrated Models," *Journal of Political Economy*, **101**, 1011-1041.
- Whelan, Karl (2000): "Balanced Growth Revisited: A Two-Sector Model of Economic Growth," *Manuscript*, Board of Governors.

## A The Data

All data are obtained from Haver Analytics (Haver mnemonics are in italics). Real output, consumption of nondurables and services, and investment (defined as gross private domestic investment plus consumption of durables) are obtained by dividing the nominal series (*GDP*, *C - CD*, and *I + CD*, respectively) by population 16 years and older (*LN16N*), and deflating using the chained-price GDP deflator (*JGDP*). Our measure of hours worked is computed by taking total hours worked reported in the National Income and Product Accounts (NIPA), which is at annual frequency. We interpolate the annual observations using growth rates computed from hours of all persons in the non-farm business sector (*LXNFH*). We divide hours worked by *LN16N* to convert them into per capita terms. Our broad measure of hours worked is consistent with our definition of output in the economy. All growth rates are computed using quarter-to-quarter log differences and then multiplied by 100 to convert them into percentages. Our data set ranges from QIII:1954 to QIV:2005. Growth rates are computed starting from QIV:1954, and we use the first four observations to initialize the lags of the VAR. Hence, the estimation sample ranges effectively from QIV:1955 to QIV:2005.

## B The Model

The following transformation induces stationarity:

$$\begin{aligned} c_t &= \frac{C_t}{Z_t}, \quad y_t = \frac{Y_t}{Z_t}, \quad i_t = \frac{I_t}{Z_t}, \quad k_t = \frac{K_t}{Z_t}, \quad \bar{k}_t = \frac{\bar{K}_t}{Z_t}, \\ w_t &= \frac{W_t}{Z_t}, \quad \xi_t = \Xi_t Z_t, \quad \xi_t^k = \Xi_t^k Z_t, \quad z_t^* = \ln(Z_t/Z_{t-1}), \end{aligned} \quad (46)$$

In terms of the detrended variables, the steady states are as follows (we take  $L_*$  as given and solve for the implied structural parameter  $\phi$ ). Return on capital:

$$r_*^k = \beta^{-1} e^\gamma - (1 - \delta). \quad (47)$$

Wages:

$$w_* = \frac{1}{1 + \lambda_f} \alpha^\alpha (1 - \alpha)^{(1-\alpha)} r_*^k^{-\alpha} L_*^{\frac{1}{1-\alpha}} \quad (48)$$

Capital stock:

$$k_* = \frac{\alpha}{1 - \alpha} \frac{w_*}{r_*^k} L_*. \quad (49)$$

Output:

$$y_* = k_*^\alpha L_*^{1-\alpha} - \Phi. \quad (50)$$

Physical capital and investment

$$\bar{k}_* = e^\gamma k_*, \quad i_* = 1 - (1 - \delta) e^{-\gamma} \bar{k}_*. \quad (51)$$

Consumption:

$$c_* = \frac{y_*}{g_*} - i_*. \quad (52)$$

Marginal utility of consumption:

$$\xi_*^k = \xi_* = c_*^{-1}(e^{z_*^*} - h)^{-1}(e^{z_*^*} - h\beta), \quad \beta = \frac{1}{r_*}e^\gamma \quad (53)$$

Labor supply:

$$\phi = \frac{w_*\xi_*}{(1 + \lambda_w)L_*^{\nu_l}}. \quad (54)$$

We conduct a first-order (log-linear) approximation of the model dynamics around the steady-state in terms of the detrended variables. Marginal product of capital:

$$\hat{r}_t^k = \hat{y}_t - \hat{K}_t. \quad (55)$$

Marginal product of labor

$$\hat{w}_t = \hat{y}_t - \hat{L}_t + \frac{2\Phi}{1 - \alpha} \left[ \beta e^{-z^*} \mathbb{E}_t[\hat{L}_{t+1}] - (1 + \beta e^{-z^*})\hat{L}_t + \hat{L}_{t-1} \right]. \quad (56)$$

Marginal utility of consumption:

$$\begin{aligned} (e^{z^*} - h\beta)(e^{z^*} - h)\hat{\xi}_t &= -(e^{2z^*} + \beta h^2)\hat{c}_t + h e^{z^*}\hat{c}_{t-1} - h e^{z^*}\hat{z}_t \\ &\quad + \beta h e^{z^*} \mathbb{E}_t[\hat{c}_{t+1}] + \beta h e^{z^*} \mathbb{E}_t[\hat{z}_{t+1}]. \end{aligned} \quad (57)$$

Capital utilization:

$$\hat{k}_t = \hat{u}_t - z_t^* + \hat{k}_{t-1}. \quad (58)$$

Capital accumulation:

$$\hat{k}_t = -\left(1 - \frac{i_*}{k_*}\right)\hat{z}_t + \left(1 - \frac{i_*}{k_*}\right)\hat{k}_{t-1} + \frac{i_*}{k_*}\mu_t + \frac{i_*}{k_*}\hat{i}_t. \quad (59)$$

Investment:

$$\frac{1}{S''e^{2z^*}}\hat{\xi}_t^k + \frac{1}{S''e^{2z^*}}\hat{\mu}_t - \frac{1}{S''e^{2z^*}}\hat{\xi}_t = \hat{z}_t - \hat{i}_{t-1} + (1 + \beta)\hat{i}_t - \beta \mathbb{E}[\hat{z}_{t+1}] - \beta \mathbb{E}[\hat{i}_{t+1}]. \quad (60)$$

Consumption Euler equation:

$$\begin{aligned} \hat{\xi}_t^k &= -\mathbb{E}_t[\hat{z}_{t+1}] + \frac{r_*^k}{r_*^k + (1 - \delta)} \mathbb{E}_t[\xi_{t+1}^k] \\ &\quad + \frac{r_*^k}{r_*^k + (1 - \delta)} \mathbb{E}_t[r_{t+1}^k] + \frac{1 - \delta}{r_*^k + (1 - \delta)} \mathbb{E}_t[\xi_{t+1}^k]. \end{aligned} \quad (61)$$

Utilization and return on capital:

$$r_*^k \hat{r}_t^k = a'' u_t. \quad (62)$$

Labor supply:

$$\hat{w}_t = \hat{\phi}_t + \nu_l \hat{L}_t - \hat{\xi}_t \quad (63)$$

Resource constraint:

$$\hat{y}_t = \hat{g}_t + \frac{c_*}{c_* + i_*} \hat{c}_t + \frac{i_*}{c_* + i_*} \hat{i}_t + \frac{r_*^k k_*}{c_* + i_*} \hat{u}_t. \quad (64)$$

Aggregate production function:

$$\hat{y}_t = \alpha \hat{k}_t + (1 - \alpha) \hat{L}_t \quad (65)$$

This system of linear rational expectations difference equations can be solved using, for instance, Sims' (2002) method. We re-normalize the investment-specific technology shock as follows:

$$\tilde{\mu}_t = \frac{1}{(1 + \beta)e^{2z^*} S^{\sigma}} \hat{\mu}_t.$$

## C Derivations

### C.1 Frequency Domain Likelihood Function

We begin by defining the  $T^* \times T^*$  unitary matrix  $W$  with elements

$$W_{j,t} = \frac{1}{\sqrt{T}} e^{i\omega_j t}$$

It can be verified that  $W^\dagger W = WW^\dagger = I_{T^*}$ . We use  $\dagger$  to denote the complex conjugate of the transpose of a complex matrix. We define the finite fourier transform  $\tilde{Y} = WY$ . The sample periodogram of  $Y$  can be expressed as

$$F_{YY}(\omega_j) = \frac{1}{2\pi} \tilde{Y}_{.j}^\dagger \tilde{Y}_{j.},$$

where  $\tilde{Y}_{.j}^\dagger$  is the  $j$ 'th column of the matrix  $\tilde{Y}^\dagger$  and  $\tilde{Y}_{j.}$  is the  $j$ 'th row of  $\tilde{Y}$ .

We write the VAR(p) as

$$Y = X\Phi + ZB + U, \quad (66)$$

where the matrix  $X$  contains the lagged  $y_t$ 's,  $Z$  contains deterministic regressors such as intercepts and time trends, and  $U$  is the  $T^* \times n$  matrix of reduced form disturbances in the VAR. According to our assumptions

$$vec(U) \sim \mathcal{N}(0, \Sigma \otimes I_{T^*}).$$

Let  $\tilde{U} = WU$  and notice that

$$vec(\tilde{U}) = (I_n \otimes W)vec(U) \sim \mathcal{N}(0, \Sigma \otimes WW^\dagger)$$

Since  $WW^\dagger = I_{T^*}$  the joint distributions of  $U$  and  $\tilde{U}$  are the same and the likelihood function for  $\tilde{U}$  is given by

$$p(\tilde{U}|\Sigma) = (2\pi)^{-nT^*/2} |\Sigma|^{-T^*/2} \exp \left\{ -\frac{1}{2} tr[\Sigma^{-1} \tilde{U}^\dagger \tilde{U}] \right\}. \quad (67)$$

We will now apply the fourier transform to (66) to obtain a relationship between  $\tilde{U}$  and  $\tilde{Y}$ :

$$\tilde{U}_{j.} = \tilde{Y}_{j.} - \tilde{X}_{j.}\Phi - \tilde{Z}_{j.}B.$$

Now let us analyze  $\tilde{X}_j$ :

$$\begin{aligned}
\tilde{X}_j &= \frac{1}{\sqrt{T^*}} \sum_{t=1}^{T^*} e^{i\omega_j t} [y'_{t-1}, \dots, y'_{t-p}] \\
&= \left[ \frac{1}{\sqrt{T^*}} \sum_{t=1}^{T^*} e^{i\omega_j t} y'_{t-1}, \dots, \frac{1}{\sqrt{T^*}} \sum_{t=1}^{T^*} e^{i\omega_j t} y'_{t-p} \right] \\
&= \left[ \frac{1}{\sqrt{T^*}} \sum_{t=1}^{T^*} e^{i\omega_j(t+1)} y'_t + \frac{1}{\sqrt{T^*}} e^{i\omega_j} y'_0 - \frac{1}{\sqrt{T^*}} e^{i\omega_j(T+1)} y'_T, \dots, \right. \\
&\quad \left. \frac{1}{\sqrt{T^*}} \sum_{t=1}^{T^*} e^{i\omega_j(t+p)} y'_t + \frac{1}{\sqrt{T^*}} \sum_{l=1}^p e^{i\omega_j} y'_{1-l} - \frac{1}{\sqrt{T^*}} \sum_{l=1}^p e^{i\omega_j(T+l)} y'_{T+l-p} \right] \\
&= \frac{1}{\sqrt{T^*}} \sum_{t=1}^{T^*} e^{i\omega_j t} y'_t [I_n e^{i\omega_j}, \dots, I_n e^{i\omega_j p}] + \text{small terms} \\
&= \tilde{Y}_j M(e^{i\omega_j}) + \text{small terms}
\end{aligned}$$

Thus, we obtain the approximation

$$\tilde{U}_j \approx \tilde{Y}_j (I_n - M(e^{i\omega_j}) \Phi) - \tilde{Z}_j B$$

and can write

$$\begin{aligned}
p(\tilde{U}|\Sigma) &\approx (2\pi)^{-nT^*/2} |\Sigma|^{-T^*/2} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \sum_{j=0}^{T^*-1} \Sigma^{-1} (\tilde{Y}_j (I_n - M(e^{i\omega_j}) \Phi) - \tilde{Z}_j B)^\dagger (\tilde{Y}_j (I_n - M(e^{i\omega_j}) \Phi) - \tilde{Z}_j B) \right] \right\} \quad (68) \\
&= (2\pi)^{-nT^*/2} |\Sigma|^{-T^*/2} \exp \left\{ -\frac{2\pi}{2} \text{tr} \sum_{j=0}^{T^*-1} (I_n - M(e^{i\omega_j}) \Phi) \Sigma^{-1} (I_n - \Phi' M'(e^{-i\omega_j})) F_{YY}(\omega_j) \right. \\
&\quad \left. + \frac{2\pi}{2} \text{tr} \sum_{j=0}^{T^*-1} B \Sigma^{-1} B' F_{ZZ}(\omega_j) + 2\pi \text{tr} \sum_{j=0}^{T^*-1} B \Sigma^{-1} (I_n - \Phi' M'(e^{-i\omega_j})) F_{YZ}(\omega_j) \right\}
\end{aligned}$$

Taking into account the Jacobian of the transformation from  $\tilde{U}$  to  $\tilde{Y}$  we obtain

$$\begin{aligned}
p(\tilde{Y}|\Phi, \Sigma) &\approx (2\pi)^{-nT^*/2} |\Sigma|^{-T^*/2} \prod_{j=0}^{T^*-1} |I_n - M(e^{i\omega_j}) \Phi| \\
&\quad \exp \left\{ -\frac{2\pi}{2} \text{tr} \sum_{j=0}^{T^*-1} (I_n - M(e^{i\omega_j}) \Phi) \Sigma^{-1} (I_n - \Phi' M'(e^{-i\omega_j})) F_{YY}(\omega_j) \right. \\
&\quad \left. + \frac{2\pi}{2} \text{tr} \sum_{j=0}^{T^*-1} B \Sigma^{-1} B' F_{ZZ}(\omega_j) + 2\pi \text{tr} \sum_{j=0}^{T^*-1} B \Sigma^{-1} (I_n - \Phi' M'(e^{-i\omega_j})) F_{YZ}(\omega_j) \right\} .
\end{aligned}$$

Finally, in the absence of deterministic trend components and using

$$S_V^{-1}(\omega_j, \Phi, \Sigma) = 2\pi (I_n - M(e^{i\omega_j}) \Phi) \Sigma^{-1} (I_n - \Phi' M'(e^{-i\omega_j}))$$

and the fact that the Jacobian of the transformation from  $\tilde{Y}$  to  $Y$  is one, we obtain

$$p(\tilde{Y}|\Phi, \Sigma) \propto \prod_{j=0}^{T^*-1} |2\pi S_V^{-1}(\omega_j, \Phi, \Sigma)|^{1/2} \exp \left\{ -\frac{1}{2} \sum_{j=0}^{T^*-1} \text{tr} [S_V^{-1}(\omega_j, \Phi, \Sigma) F_{YY}(\omega_j)] \right\} .$$

## C.2 No Adjustment under Equal Weights

Let  $\omega_j = 2\pi j/m$  for  $j = 0, \dots, m-1$ . We express the integral of interest as Riemann sum

$$\int_0^{2\pi} \ln \frac{1}{2\pi} S_V^{-1}(\omega, \Phi, \Sigma) d\omega = \lim_{m \rightarrow \infty} \frac{2\pi}{m} \sum_{j=0}^{m-1} \ln \frac{1}{2\pi} S_V^{-1}(\omega_j, \Phi, \Sigma)$$

and will study the right-hand-side limit. The subsequent calculations are conducted for a VAR(1). They can be easily generalized by re-writing a VAR(p) in companion form.

The calculation is based on an argument by Espasa (1977) as reproduced in Engle (1980). We write the VAR as

$$y_t = \Phi_1 y_{t-1} + u_t. \quad (69)$$

The system can be transformed through a complex Schur decomposition of  $\Phi_1$ . There exist matrices  $Q$  and  $\Lambda$  such that  $Q\Lambda Q' = \Phi_1$ ,  $Q'Q = QQ' = I$ , and  $\Lambda$  is uppertriangular. Moreover, Let  $x_t = Qy_t$  and premultiply the above equation by  $Q$  to obtain:

$$x_t = \Lambda x_{t-1} + Qu_t \quad (70)$$

Since  $y_t = Q'x_t$  we deduce that

$$S_V(\omega_j, \Phi, \Sigma) = Q' S_V^x(\omega_j, \Lambda, Q\Sigma Q') Q,$$

where  $S_V^x(\cdot)$  denotes the spectral density matrix of the transformed endogenous variables  $x_t$ . Hence,

$$\begin{aligned} S_V^{-1}(\omega, \Phi, \Sigma) &= Q [S_V^x(\omega, \Lambda, Q\Sigma Q')]^{-1} Q' \\ &= 2\pi Q [I - \Lambda e^{i\omega}] Q' \Sigma^{-1} Q [I - \Lambda' e^{-i\omega}] Q' \end{aligned}$$

and

$$\begin{aligned} \frac{1}{m} \sum_{j=0}^{m-1} \ln \frac{1}{2\pi} S_V^{-1}(\omega_j, \Phi, \Sigma) &= \frac{1}{m} \sum_{j=0}^{m-1} \ln Q [I - \Lambda e^{i\omega_j}] Q' \Sigma^{-1} Q [I - \Lambda' e^{-i\omega_j}] Q' \\ &= \frac{1}{m} \sum_{j=0}^{m-1} -\ln |\Sigma| + 2 \ln |I - \Lambda e^{i\omega_j}| \\ &= -\ln |\Sigma| + \frac{2}{m} \sum_{l=1}^n \ln \prod_{j=0}^{m-1} |1 - \Lambda_{ll} e^{i\omega_j}| \end{aligned}$$

where  $\Lambda_{ll}$  is the  $l$ 'th diagonal term of  $\Lambda$ . Now consider the second term. Notice that

$$\prod_{j=0}^{m-1} (X - e^{i\omega_j}) = X^m - 1.$$

Therefore, as  $m \rightarrow \infty$

$$\begin{aligned} \sum_{l=1}^n \ln \prod_{j=0}^{m-1} (1 - \Lambda_{ll} e^{i\omega_j}) &= \sum_{l=1}^n \ln \Lambda_{ll}^m \prod_{j=0}^{m-1} (1/\Lambda_{ll} - e^{i\omega_j}) \\ &= \sum_{l=1}^n \ln(1 - \Lambda_{ll}^m) \rightarrow 0 \end{aligned}$$



and we deduce that

$$\frac{1}{m} \sum_{j=0}^{m-1} \ln S_V^{-1}(\omega_j, \Phi, \Sigma) \longrightarrow -\ln |\Sigma|$$

as long as the eigenvalues of  $\Lambda_U$  of the matrix  $\Phi_1$  are less than one in absolute value.  $\square$

### C.3 Quadratic Expansion of Adjustment Term

We begin by presenting two Lemmas that will be helpful for the subsequent analysis. Define the symmetric  $n^2 \times n^2$  matrix  $D$  as

$$D = [I_n \otimes \iota_1, \dots, I_n \otimes \iota_n]$$

where  $\iota_j$  is a  $j \times 1$  unit vector with the  $j$ 'th element equal to one.

**Lemma 1** *Let  $A$  be a  $n \times k$  real matrix and  $B$  be a  $k \times n$  real matrix. Then*

$$\text{tr}[ABAB] = \text{vec}(B)'(I_n \otimes A')D(I_n \otimes A)\text{vec}(B) \quad \square$$

**Proof of Lemma 1:** Notice that  $\text{vec}((AB)') = D\text{vec}(AB)$ . It can be verified by direct matrix multiplication that

$$\text{tr}[ABAB] = [\text{vec}((AB)')]'\text{vec}(AB).$$

Hence, we obtain the desired result:

$$\begin{aligned} \text{tr}[ABAB] &= [\text{vec}(AB)]'D\text{vec}(AB) \\ &= \text{vec}(B)'(I_n \otimes A')D(I_n \otimes A)\text{vec}(B). \quad \square \end{aligned}$$

**Lemma 2** *Let  $C = A + iB$  be a  $n \times n$  complex matrix. Then*

$$\text{tr}[CC] + \text{tr}[C^\dagger C^\dagger] = 2\text{tr}[AA] - 2\text{tr}[BB]. \quad \square$$

**Proof of Lemma 2:** follows from direct matrix manipulations:

$$\begin{aligned} \text{tr}[CC] + \text{tr}[C^\dagger C^\dagger] &= \text{tr}[(A + iB)(A + iB)] + \text{tr}[(A' - iB')(A' - iB')] \\ &= \text{tr}[AA] + 2i\text{tr}[AB] - \text{tr}[BB] + \text{tr}[A'A'] - 2i\text{tr}[A'B'] - \text{tr}[B'B'] \\ &= 2\text{tr}[AA] - 2\text{tr}[BB]. \quad \square \end{aligned}$$

We now proceed with an expansion of the term

$$\ln |S_V^{-1}(\omega_j, \Phi, \Sigma)|$$

around  $\Phi = \tilde{\Phi}$ . First, we will take derivatives of  $S_V^{-1}(\omega_j, \Phi, \Sigma)$  with respect to  $\Phi$ :

$$\begin{aligned} dS_V^{-1}(\omega_j, \Phi, \Sigma) &= -2\pi M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j})) - 2\pi(I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j}) \\ d^2S_V^{-1}(\omega_j, \Phi, \Sigma) &= 4\pi M(e^{i\omega_j})d\Phi\Sigma^{-1}d\Phi'M'(e^{-i\omega_j}) \end{aligned}$$

Second, we take derivatives of  $\ln |S_V^{-1}(\omega_j, \Phi, \Sigma)|$  with respect to  $S_V^{-1}(\omega_j, \Phi, \Sigma)$ :

$$\begin{aligned} d \ln |S_V^{-1}(\omega_j, \Phi, \Sigma)| &= \text{tr}[S_V(\omega_j, \Phi, \Sigma) dS_V^{-1}(\omega_j, \Phi, \Sigma)] \\ d^2 \ln |S_V^{-1}(\omega_j, \Phi, \Sigma)| &= -\text{tr}[S_V(\omega_j, \Phi, \Sigma) dS_V^{-1}(\omega_j, \Phi, \Sigma) S_V(\omega_j, \Phi, \Sigma) dS_V^{-1}(\omega_j, \Phi, \Sigma)]. \end{aligned}$$

Define  $d\Phi = \Phi - \tilde{\Phi}$ . Hence, we obtain

$$\begin{aligned} \ln |S_V^{-1}(\omega_j, \Phi, \Sigma)| &= \ln |S_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)| \\ &\quad -\text{tr} \ 2\pi\Sigma^{-1}(I_n - \tilde{\Phi}'M'(e^{-i\omega_j}))S_V(\omega_j, \tilde{\Phi}, \Sigma)M(e^{i\omega_j})d\Phi \\ &\quad -\text{tr} \ 2\pi\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \tilde{\Phi}, \Sigma)(I_n - M(e^{i\omega_j})\tilde{\Phi}) \\ &\quad +\frac{1}{2}\text{tr} \ 4\pi\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \tilde{\Phi}, \Sigma)M(e^{i\omega_j})d\Phi \\ &\quad -\frac{1}{2}\text{tr} \ S_V(\omega_j, \tilde{\Phi}, \Sigma)dS_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)S_V(\omega_j, \tilde{\Phi}, \Sigma)dS_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma) \\ &\quad +\text{small} \\ &= \ln |S_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)| \\ &\quad -\text{tr} \ (I_n - M(e^{i\omega_j})\tilde{\Phi})^{-1}M(e^{i\omega_j})d\Phi \ -\text{tr} \ d\Phi'M'(e^{-i\omega_j})(I_n - \tilde{\Phi}'M'(e^{-i\omega_j}))^{-1} \\ &\quad +\frac{1}{2}\text{tr} \ 4\pi\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \tilde{\Phi}, \Sigma)M(e^{i\omega_j})d\Phi \\ &\quad -\frac{1}{2}\text{tr} \ S_V(\omega_j, \tilde{\Phi}, \Sigma)dS_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)S_V(\omega_j, \tilde{\Phi}, \Sigma)dS_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma) \\ &\quad +\text{small}. \end{aligned}$$

Now consider the last term (omitting tildes):

$$\begin{aligned} &\text{tr} \ S_V(\omega_j, \Phi, \Sigma)dS_V^{-1}(\omega_j, \Phi, \Sigma)S_V(\omega_j, \Phi, \Sigma)dS_V^{-1}(\omega_j, \Phi, \Sigma) \\ &= \text{tr} \ (2\pi)^2 S_V(\omega_j, \Phi, \Sigma) \ M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j})) + (I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j}) \\ &\quad \times S_V(\omega_j, \Phi, \Sigma) \ M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j})) + (I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j}) \\ &= \text{tr} \ (2\pi)^2 S_V(\omega_j, \Phi, \Sigma)M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j}))S_V(\omega_j, \Phi, \Sigma)M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j})) \\ &\quad +\text{tr} \ (2\pi)^2 S_V(\omega_j, \Phi, \Sigma)M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j}))S_V(\omega_j, \Phi, \Sigma)(I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j}) \\ &\quad +\text{tr} \ (2\pi)^2 S_V(\omega_j, \Phi, \Sigma)(I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \Phi, \Sigma)M(e^{i\omega_j})d\Phi\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j})) \\ &\quad +\text{tr} \ (2\pi)^2 S_V(\omega_j, \Phi, \Sigma)(I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \Phi, \Sigma)(I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}d\Phi'M'(e^{-i\omega_j}) \\ &= \text{tr} \ (I_n - M(e^{i\omega_j})\Phi)^{-1}M(e^{i\omega_j})d\Phi(I_n - M(e^{i\omega_j})\Phi)^{-1}M(e^{i\omega_j})d\Phi \\ &\quad +\text{tr} \ d\Phi'M'(e^{-i\omega_j})(I_n - \Phi'M'(e^{-i\omega_j}))^{-1}d\Phi'M'(e^{-i\omega_j})(I_n - \Phi'M'(e^{-i\omega_j}))^{-1} \\ &\quad +2\text{tr} \ 2\pi\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \Phi, \Sigma)M(e^{i\omega_j})d\Phi \ . \end{aligned}$$

We will focus on the first two terms in this expression. Notice that  $d\Phi$  is a real  $k \times n$  matrix, whereas

$$F(\omega_j, \Phi) = (I_n - M(e^{i\omega_j})\Phi)^{-1}M(e^{i\omega_j})$$

is a  $n \times k$  complex matrix. Let  $C = \text{re}(F(\omega_j, \Phi))d\Phi + i\text{im}(F(\omega_j, \Phi))d\Phi$  and apply Lemmas 1 and 2. Define  $d\phi = \text{vec}(d\Phi)$ . Hence,

$$\begin{aligned} & \text{tr } S_V(\omega_j, \Phi, \Sigma)dS_V^{-1}(\omega_j, \Phi, \Sigma)S_V(\omega_j, \Phi, \Sigma)dS_V^{-1}(\omega_j, \Phi, \Sigma) \\ &= 2d\phi' \begin{bmatrix} I_n \otimes \text{re}(F(\omega_j, \Phi))' & D \\ I_n \otimes \text{re}(F(\omega_j, \Phi)) & -I_n \otimes \text{im}(F(\omega_j, \Phi))' \\ I_n \otimes \text{im}(F(\omega_j, \Phi)) & D \\ I_n \otimes \text{im}(F(\omega_j, \Phi)) & -I_n \otimes \text{re}(F(\omega_j, \Phi))' \end{bmatrix} d\phi \\ & \quad + 2\text{tr } 2\pi\Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_V(\omega_j, \Phi, \Sigma)M(e^{i\omega_j})d\Phi \end{aligned}$$

Combining terms and using the definition of  $F(\omega_j, \Phi)$ , we obtain the desired quadratic expansion:

$$\begin{aligned} & \ln |S_V^{-1}(\omega_j, \Phi, \Sigma)| \\ &= \ln |S_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)| - \text{tr } F(\omega_j, \tilde{\Phi})d\Phi - \text{tr } d\Phi'F^\dagger(\omega_j, \tilde{\Phi}) \\ & \quad - d\phi' \begin{bmatrix} I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi}))' & D \\ I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi})) & -I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi}))' \\ I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi})) & D \\ I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi})) & -I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi}))' \end{bmatrix} d\phi \\ & \quad + \text{small} \\ &= \ln |S_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)| - 2\text{vec } \text{re}(F'(\omega_j, \tilde{\Phi}))' d\phi \\ & \quad - d\phi' \begin{bmatrix} I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi}))' & D \\ I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi})) & -I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi}))' \\ I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi})) & D \\ I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi})) & -I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi}))' \end{bmatrix} d\phi \\ & \quad + \text{small}. \end{aligned}$$

## C.4 Gaussian Approximation of the Conditional Prior of $\Phi$

We proceed with a quadratic approximation of the “regular” exponential term in the frequency domain likelihood function around  $\tilde{\Phi}$  :

$$\begin{aligned}
& \frac{1}{2\pi} \text{tr}[S_V^{-1}(\omega_j, \Phi, \Sigma)S_D(\omega_j, \theta)] \\
&= \text{tr} (I_n - M(e^{i\omega_j})\Phi)\Sigma^{-1}(I_n - \Phi'M'(e^{-i\omega_j}))S_D(\omega_j, \theta) \\
&= \text{tr} \begin{bmatrix} I_n - M(e^{i\omega_j})\tilde{\Phi} - M(e^{i\omega_j})d\Phi & \Sigma^{-1} \\ I_n - \tilde{\Phi}'M'(e^{-i\omega_j}) - d\Phi'M'(e^{-i\omega_j}) & S_D(\omega_j, \theta) \end{bmatrix} \\
&= \text{tr} \Sigma^{-1}(I_n - \tilde{\Phi}'M'(e^{-i\omega_j}))S_D(\omega_j, \theta)(I_n - M(e^{i\omega_j})\tilde{\Phi}) \\
&\quad - \text{tr} \Sigma^{-1}(I_n - \tilde{\Phi}'M'(e^{-i\omega_j}))S_D(\omega_j, \theta)M(e^{i\omega_j})d\Phi \\
&\quad - \text{tr} \Sigma^{-1}d\Phi'M'(e^{-i\omega_j})S_D(\omega_j, \theta)(I_n - M(e^{i\omega_j})\tilde{\Phi}) \\
&\quad + d\phi' \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) d\phi \\
&= \text{tr} \Sigma^{-1}(I_n - \tilde{\Phi}'M'(e^{-i\omega_j}))S_D(\omega_j, \theta)(I_n - M(e^{i\omega_j})\tilde{\Phi}) \\
&\quad - 2\text{vec re}(M'(e^{-i\omega_j}))S_D(\omega_j, \theta)' \Sigma^{-1} \otimes I_k d\phi + 2\tilde{\phi}' \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) d\phi \\
&\quad + d\phi' \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) d\phi.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& -\frac{1}{2\pi} \ln |S_V^{-1}(\omega_j, \Phi, \Sigma)| + \frac{1}{2\pi} \text{tr}[S_V^{-1}(\omega_j, \Phi, \Sigma)S_D(\omega_j, \theta)] \\
&= -\frac{1}{2\pi} \ln |S_V^{-1}(\omega_j, \tilde{\Phi}, \Sigma)| + \text{tr} \Sigma^{-1}(I_n - \tilde{\Phi}'M'(e^{-i\omega_j}))S_D(\omega_j, \theta)(I_n - M(e^{i\omega_j})\tilde{\Phi}) \\
&\quad + 2\frac{1}{2\pi} \text{vec re}(F'(\omega_j, \tilde{\Phi}))' I_n \otimes I_k d\phi \\
&\quad - 2\text{vec re}(M'(e^{-i\omega_j}))S_D(\omega_j, \theta)' \Sigma^{-1} \otimes I_k d\phi + 2\tilde{\phi}' \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) d\phi \\
&\quad + \frac{1}{2\pi} d\phi' \begin{bmatrix} I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi}))' & D \\ I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi})) & - \end{bmatrix} \begin{bmatrix} I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi}))' & D \\ I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi})) & - \end{bmatrix} d\phi \\
&\quad + d\phi' \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) d\phi.
\end{aligned}$$

Now define

$$\begin{aligned}
V^{-1}(\omega_j, \tilde{\Phi}, \Sigma, \theta) &= \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) \\
&\quad + \frac{1}{2\pi} \begin{bmatrix} I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi}))' & D \\ I_n \otimes \text{re}(F(\omega_j, \tilde{\Phi})) & - \end{bmatrix} \\
&\quad - \frac{1}{2\pi} \begin{bmatrix} I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi}))' & D \\ I_n \otimes \text{im}(F(\omega_j, \tilde{\Phi})) & - \end{bmatrix}
\end{aligned}$$

and

$$\begin{aligned} \mu(\omega_j, \tilde{\Phi}, \Sigma, \theta) &= \Sigma^{-1} \otimes I_k \text{vec re}(M'(e^{-i\omega_j}))S_D(\omega_j, \theta) - \Sigma^{-1} \otimes M'(e^{-i\omega_j})S_D(\omega_j, \theta)M(e^{i\omega_j}) \tilde{\phi} \\ &\quad - \frac{1}{2\pi} I_n \otimes I_k \text{vec re}(F'(\omega_j, \tilde{\Phi})) \end{aligned}$$

Hence,

$$\begin{aligned} \int \lambda(\omega)V^{-1}(\omega, \tilde{\Phi}, \Sigma, \theta)d\omega &= \Sigma^{-1} \otimes \Gamma_{XX, \lambda}(\theta) \\ &\quad + \frac{1}{2\pi} \int \lambda(\omega) I_n \otimes \text{re}(F(\omega, \tilde{\Phi}))' D I_n \otimes \text{re}(F(\omega, \tilde{\Phi})) d\omega \\ &\quad - \frac{1}{2\pi} \int \lambda(\omega) I_n \otimes \text{im}(F(\omega, \tilde{\Phi}))' D I_n \otimes \text{im}(F(\omega, \tilde{\Phi})) d\omega \end{aligned}$$

and

$$\begin{aligned} \int \lambda(\omega)\mu(\omega, \tilde{\Phi}, \Sigma, \theta)d\omega &= \Sigma^{-1} \otimes I_k \text{vec}[\Gamma_{XY, \lambda}(\theta)] - \Sigma^{-1} \otimes \Gamma_{XX, \lambda}(\theta) \tilde{\phi} \\ &\quad - \frac{1}{2\pi} I_n \otimes I_k \text{vec} \int \lambda(\omega)\text{re}(F'(\omega, \tilde{\Phi}))d\omega \end{aligned}$$

We can therefore deduce that the posterior of  $\Phi$  given  $\Sigma$  and  $\theta$  can be approximated by

$$\phi|\Sigma, \theta \sim \mathcal{N} \left( \tilde{\phi} + \int \lambda(\omega)V^{-1}(\omega, \tilde{\Phi}, \Sigma, \theta)d\omega^{-1} \int \lambda(\omega)\mu(\omega, \tilde{\Phi}, \Sigma, \theta)d\omega, \int \lambda(\omega)V^{-1}(\omega, \tilde{\Phi}, \Sigma, \theta)d\omega^{-1} \right).$$

To guarantee that the conditional prior distribution of  $\Sigma$  given  $\Phi$  belongs to the inverted Wishart family after we have replaced  $f_{\lambda, T^*}(\Phi)$  by a quadratic expansion, we must choose a  $\tilde{\Phi}$  that is independent of  $\Sigma$ , but at the same attains a high posterior density. We construct  $\tilde{\Phi}$  as follows. Recall that in the absence of approximations our prior density is of the form

$$\begin{aligned} p(\Phi, \Sigma|\theta) &\propto \mathcal{I}_{\{\Phi \in \mathcal{P}\}} |\Sigma|^{-(T^*+n+1)/2} f_{\lambda, T^*}(\Phi) \\ &\quad \times \exp \left[ -\frac{T^*}{2} \text{tr} \Sigma^{-1} \left( \Gamma_{\lambda, YY}(\theta) - 2\Gamma_{\lambda, YX}(\theta)\Phi + \Phi' \Gamma_{\lambda, XX}(\theta)\Phi \right) \right]. \end{aligned}$$

Define

$$S = T^* \left( \Gamma_{\lambda, YY}(\theta) - \Gamma_{\lambda, YX}(\theta)\Phi - \Phi' \Gamma_{\lambda, XY} + \Phi' \Gamma_{\lambda, XX}(\theta)\Phi \right)$$

and notice that the conditional density  $p(\Sigma|\Phi, \theta)$  is of the inverted Wishart form. Using the fact that an inverted Wishart distribution with parameters  $S$  and  $T^*$  has a density that is proportional to

$$|S|^{T^*/2} |\Sigma|^{-(T^*+n+1)/2} \exp \left[ -\frac{1}{2} \text{tr}[\Sigma^{-1}S] \right]$$

we deduce that

$$p(\Phi|\theta) \propto \mathcal{I}_{\{\Phi \in \mathcal{P}\}} f_{\lambda, T^*}(\Phi) \left( \Gamma_{\lambda, YY}(\theta) - \Gamma_{\lambda, YX}(\theta)\Phi - \Phi' \Gamma_{\lambda, XY} + \Phi' \Gamma_{\lambda, XX}(\theta)\Phi \right)^{-T^*/2}$$

and define

$$\tilde{\Phi} = \text{argmax}_{\Phi} p(\Phi|\theta).$$

We then replace  $\ln f_{\lambda, T^*}(\Phi)$  by a quadratic approximation around  $\tilde{\Phi}$ .

**Example:** Consider the case of the AR(1) model. The inverse spectral density is given by

$$S_V^{-1}(\omega, \phi, \sigma^2) = \frac{2\pi}{\sigma^2}(1 + \phi^2 - 2\phi \cos \omega).$$

Moreover,  $M(z) = 1$  and  $D = 1$ . It can be verified by straightforward algebraic manipulations that

$$F(\omega, \phi) = \frac{\cos \omega - \phi}{1 + \phi^2 - 2\phi \cos \omega} + i \frac{\sin \omega}{1 + \phi^2 - 2\phi \cos \omega}.$$

Hence,

$$\begin{aligned} \ln |S_V^{-1}(\omega, \phi, \sigma^2)| &\approx \ln |S_V^{-1}(\omega, \tilde{\phi}, \sigma^2)| \\ &\quad - 2 \frac{\cos(\omega) - \tilde{\phi}}{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega} (\phi - \tilde{\phi}) \\ &\quad - \frac{\tilde{\phi}^2 - 2\tilde{\phi} \cos \omega + 2 \cos^2 \omega - 1}{(1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega)^2} (\phi - \tilde{\phi})^2 \end{aligned}$$

Moreover,

$$\begin{aligned} \frac{1}{2\pi} \text{tr}[S_V^{-1}(\omega, \phi, \sigma^2) S_D(\omega)] &= \frac{S_D(\omega)}{\sigma^2} (1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega) \\ &\quad - 2 \frac{S_D(\omega)}{\sigma^2} (\cos \omega - \tilde{\phi})(\phi - \tilde{\phi}) \\ &\quad + \frac{S_D(\omega)}{\sigma^2} (\phi - \tilde{\phi})^2 \end{aligned}$$

To approximate

$$-\frac{1}{2\pi} \ln |S_V^{-1}(\omega, \phi, \sigma^2)| + \frac{1}{2\pi} \text{tr}[S_V^{-1}(\omega, \phi, \sigma^2) S_D(\omega)]$$

we define the variance and mean function

$$\begin{aligned} V^{-1}(\omega, \tilde{\phi}, \sigma^2) &= \frac{S_D(\omega)}{\sigma^2} + \frac{1}{2\pi} \frac{\tilde{\phi}^2 - 2\tilde{\phi} \cos \omega + 2 \cos^2 \omega - 1}{(1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega)^2} \\ \mu(\omega, \tilde{\phi}, \sigma^2) &= \frac{S_D(\omega)}{\sigma^2} (\cos \omega - \tilde{\phi}) - \frac{1}{2\pi} \frac{\cos \omega - \tilde{\phi}}{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega} \end{aligned}$$

Using the notation

$$\gamma_{\lambda,0} = \int_0^{2\pi} \lambda(\omega) S_D(\omega) d\omega, \quad \gamma_{\lambda,1} = \int_0^{2\pi} \lambda(\omega) \cos(\omega) S_D(\omega) d\omega$$

we can write

$$\begin{aligned} \int \lambda(\omega) V^{-1}(\omega, \tilde{\phi}, \sigma^2) d\omega &= \frac{1}{\sigma^2} \gamma_{\lambda,0} + \frac{1}{2\pi} \int \lambda(\omega) \frac{\tilde{\phi}^2 - 2\tilde{\phi} \cos \omega + 2 \cos^2 \omega - 1}{(1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega)^2} d\omega \\ \int \lambda(\omega) \mu(\omega, \tilde{\phi}, \sigma^2) d\omega &= \frac{1}{\sigma^2} (\gamma_{\lambda,1} - \tilde{\phi} \gamma_{\lambda,0}) - \frac{1}{2\pi} \int \lambda(\omega) \frac{\cos \omega - \tilde{\phi}}{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos \omega} d\omega \end{aligned}$$

## C.5 Models with Intercepts and Trends

Consider the VAR given in (37). Let  $\Psi = [\Psi_0, \Psi_1]'$ ,  $\psi = \text{vec}(\Psi')$ , and  $z_t' = [1, t]$ . Moreover, define  $\tilde{y}_t'(\Psi) = y_t' - z_t' \Psi$  and let  $\tilde{Y}(\Psi)$  be the  $T \times n$  matrix with rows  $\tilde{y}_t'(\Psi)$  and  $\tilde{X}(\Psi)$  be the  $T \times np$  matrix with rows

$$\tilde{x}_t'(\Psi) = [\tilde{y}_{t-1}'(\Psi), \dots, \tilde{y}_{t-1}'(\Psi)]'.$$

Using this notation,

$$\tilde{y}'_i(\Psi) = \tilde{x}'_i(\Psi) \cdot \Phi + u'_i.$$

Now define:

$$\hat{\Gamma}_{YY}(\Psi) = \tilde{Y}(\Psi)' \tilde{Y}(\Psi) / T, \quad \hat{\Gamma}_{YX}(\Psi) = \tilde{Y}(\Psi)' \tilde{X}(\Psi) / T, \quad \hat{\Gamma}_{XX}(\Psi) = \tilde{X}(\Psi)' \tilde{X}(\Psi) / T.$$

The likelihood function can then be written as

$$p(Y|\Psi, \Phi, \Sigma) = (2\pi)^{-\frac{nT}{2}} |\Sigma|^{-T/2} \exp \left[ -\frac{T}{2} \text{tr}[\Sigma^{-1}(\hat{\Gamma}_{YY}(\Psi) - 2\hat{\Gamma}_{YX}(\Psi)\Phi + \Phi'\hat{\Gamma}_{XX}\Phi)] \right]. \quad (71)$$

We combine the likelihood with a prior of the form

$$p(\Psi, \Phi, \Sigma|\theta) = p(\Phi, \Sigma|\theta)p(\Psi|\theta) \quad (72)$$

where

$$\begin{aligned} p(\Phi, \Sigma|\theta) &\propto \mathcal{I}_{\{\Phi \in \mathcal{P}\}} |\Sigma|^{-(T^*+n+1)/2} f_{\lambda, T^*}(\Phi) \\ &\quad \times \exp \left[ -\frac{T^*}{2} \text{tr} \Sigma^{-1} \Gamma_{\lambda, Y Y}(\theta) - 2\Gamma_{\lambda, Y X}(\theta)\Phi + \Phi'\Gamma_{\lambda, X X}(\theta)\Phi \right] \\ p(\Psi|\theta) &\propto |V_0^\psi|^{-1/2} \exp \left[ -\frac{1}{2}(\psi - \mu_0^\psi(\theta))'(V_0^\psi)^{-1}(\psi - \mu_0^\psi(\theta)) \right]. \end{aligned}$$

We use the following mean vector and covariance matrix for  $\psi$ :

$$\begin{aligned} \mu_0^{\psi'} &= y_{adj}, y_{adj} + \ln(c^*/y^*), y_{adj} + \ln(i^*/y^*), h_{adj}, \gamma, \gamma, \gamma, h_{adj} \\ V_0^\psi &= \begin{bmatrix} \tau_{0,1} & \tau_{0,1} & \tau_{0,1} & 0 & 0 & 0 & 0 & 0 \\ \tau_{0,1} & \tau_{0,1} + \tau_{0,2} & \tau_{0,1} & 0 & 0 & 0 & 0 & 0 \\ \tau_{0,1} & \tau_{0,1} & \tau_{0,1} + \tau_{0,3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau_{0,4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \tau_{1,1} & \tau_{1,1} & \tau_{1,1} & 0 \\ 0 & 0 & 0 & 0 & \tau_{1,1} & \tau_{1,1} + \tau_{1,2} & \tau_{1,1} & 0 \\ 0 & 0 & 0 & 0 & \tau_{1,1} & \tau_{1,1} & \tau_{1,1} + \tau_{1,3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \tau_{0,4} \end{bmatrix}. \end{aligned}$$

In turn, we will derive the conditional posterior densities that can be used in a Gibbs sampling scheme.

Using the notation that, for instance,

$$\tilde{\Gamma}_{\lambda, \zeta, Y Y}(\theta, \Psi) = \zeta \Gamma_{\lambda, Y Y}(\theta) + (1 - \zeta) \hat{\Gamma}_{Y Y}(\Psi)$$

we define

$$\begin{aligned} \tilde{\Phi}_{\lambda, \zeta}(\theta, \Psi) &= \tilde{\Gamma}_{\lambda, \zeta, X X}^{-1}(\theta, \Psi) \tilde{\Gamma}_{\lambda, \zeta, X Y}(\theta, \Psi), \\ \tilde{\Sigma}_{\lambda, \zeta}(\theta, \Psi) &= \tilde{\Gamma}_{\lambda, \zeta, Y Y}(\theta, \Psi) - \tilde{\Gamma}_{\lambda, \zeta, Y X}(\theta, \Psi) \tilde{\Gamma}_{\lambda, \zeta, X X}^{-1}(\theta, \Psi) \tilde{\Gamma}_{\lambda, \zeta, X Y}(\theta, \Psi). \end{aligned}$$

and write the conditional posterior density as

$$p(\Phi, \Sigma | Y, \Psi \theta) \propto \mathcal{I}_{\{\Phi \in \text{int}(\mathcal{P})\}} f_{\lambda, T^*}(\Phi) \quad (73)$$

$$\times p_{\mathcal{I}W-\mathcal{N}}(\Phi, \Sigma, \tilde{\Phi}_{\lambda, \zeta}(\theta, \Psi), \tilde{\Sigma}_{\lambda, \zeta}(\theta, \Psi), \tilde{\Gamma}_{\lambda, \zeta, XX}(\theta, \Psi), T^* + T) .$$

To study the posterior density of  $\Psi$  it is convenient to rewrite the likelihood function as follows. Define  $\psi = \text{vec}(\Psi')$  and notice that the VAR can be expressed as

$$y_t - \sum_{j=1}^p \Phi_j y_{t-j} = \left( I - \sum_{j=1}^p \Phi_j \right) \Psi_0 + \left( I \cdot t - \sum_{j=1}^p \Phi_j (t-j) \right) \Psi_1 + u_t$$

or

$$\hat{y}_t = A_t \psi + u_t,$$

where

$$\hat{y}_t = y_t - \sum_{j=1}^p \Phi_j y_{t-j} \quad \text{and} \quad A_t = \left( I - \sum_{j=1}^p \Phi_j \right), \quad \left( I \cdot t - \sum_{j=1}^p \Phi_j (t-j) \right) .$$

Hence, we can express the kernel of the likelihood function as

$$\begin{aligned} & -\frac{1}{2} tr \Sigma^{-1} (\tilde{Y}(\Psi) - \tilde{X}(\Psi) \cdot \Phi)' (\tilde{Y}(\Psi) - \tilde{X}(\Psi) \cdot \Phi) \\ &= -\frac{1}{2} \sum_{t=1}^T (\hat{y}_t - A_t \psi)' \Sigma^{-1} (\hat{y}_t - A_t \psi) \\ &= -\frac{1}{2} \left[ \sum_{t=1}^T \hat{y}_t' \Sigma^{-1} \hat{y}_t - 2 \sum_{t=1}^T \hat{y}_t' \Sigma^{-1} A_t \right] \psi + \psi' \left[ \sum_{t=1}^T A_t' \Sigma^{-1} A_t \right] \psi . \end{aligned}$$

We deduce that

$$\psi | Y, \Phi, \Sigma, \theta \sim \mathcal{N}(\mu_T^\psi, V_T^\psi), \quad (74)$$

where

$$\begin{aligned} V_T^\psi &= \left( (V_0^\psi)^{-1} + \sum_{t=1}^T A_t' \Sigma^{-1} A_t \right)^{-1} \\ \mu_T^\psi &= V_T^\psi \left( (V_0^\psi)^{-1} \mu_0^\psi + \sum_{t=1}^T \hat{y}_t' \Sigma^{-1} A_t \right)' . \end{aligned}$$

## D Computational Issues

**Computation of Adjustment Term.** Let  $\Lambda_{ll}$ ,  $l = 1, \dots, np$  be the possibly complex eigenvalues of the matrix of autoregressive coefficients for the VAR(p) (written in companion form). We approximate the log adjustment term as follows:

$$\begin{aligned} \ln f_{\lambda, T^*}(\Phi) &= \frac{T^*}{2 \cdot 2\pi} \int_0^{2\pi} \lambda(\omega) \ln |(I - \Phi' M'(e^{i\omega}))(I - M(e^{-i\omega}) \Phi)| d\omega \\ &\approx \frac{T^*}{2} \sum_{l=1}^{np} \frac{2}{m} \sum_{j=0}^{m-1} \lambda(\omega_j) \ln |1 - \Lambda_{ll} e^{i\omega_j}| \\ &= \frac{T^*}{2} \sum_{l=1}^{np} \frac{1}{m} \sum_{j=0}^{m-1} \lambda(\omega_j) \ln(1 + |\Lambda_{ll}|^2 - 2 \text{re}(\Lambda_{ll}) \cos(\omega_j)) - 2 \text{im}(\Lambda_{ll}) \sin(\omega_j) \end{aligned}$$



Table 1: DSGE Model's Parameter Estimates

	Domain	Distr.	Prior			Posterior (I)		Posterior (II)	
			P(1)	P(2)	Interval	Mean	Interval	Mean	Interval
$\alpha$	$[0, 1)$	Beta	0.33	0.10	[0.17, 0.49]	0.23	[0.21, 0.27]	0.27	[0.26, 0.29]
$\Phi$	$\mathbb{R}^+$	Gamma	33.00	15.00	[9.51, 55.40]	5.88	[3.20, 8.65]	30.50	[19.85, 42.86]
$s'$	$\mathbb{R}^+$	Gamma	4.00	1.50	[1.61, 6.31]	1.30	[0.51, 2.02]	0.98	[0.39, 1.58]
$h$	$[0, 1)$	Beta	0.70	0.05	[0.62, 0.78]	0.78	[0.73, 0.82]	0.79	[0.74, 0.84]
$a''$	$\mathbb{R}^+$	Gamma	0.20	0.10	[0.05, 0.35]	0.31	[0.14, 0.46]	0.28	[0.12, 0.44]
$\nu_l$	$\mathbb{R}^+$	Gamma	2.00	0.75	[0.81, 3.16]	3.68	[2.40, 4.92]	3.17	[1.44, 4.93]
$\gamma$	$\mathbb{R}^+$	Gamma	2.00	1.00	[0.48, 3.49]	1.06	[0.62, 1.51]	1.47	[0.99, 1.94]
$g^*$	$[0, 1)$	Beta	0.30	0.10	[0.14, 0.46]	0.18	[0.08, 0.26]	0.24	[0.23, 0.25]
$L_{adj}$	$\mathbb{R}$	Normal	252	10.0	[235, 269]	248	[235, 261]	251	[242, 261]
$\rho_\phi$	$[0, 1)$	Beta	0.90	0.05	[0.83, 0.98]	0.97	[0.95, 1.00]	0.90	fixed
$\rho_\mu$	$[0, 1)$	Beta	0.90	0.05	[0.83, 0.98]	0.97	[0.95, 1.00]	0.90	fixed
$\rho_g$	$[0, 1)$	Beta	0.90	0.05	[0.83, 0.98]	0.99	[0.99, 1.00]	0.90	fixed
$\sigma_z$	$\mathbb{R}^+$	InvGamma	0.75	2.00	[0.31, 2.35]	1.09	[1.00, 1.19]	1.14	[1.04, 1.24]
$\sigma_\phi$	$\mathbb{R}^+$	InvGamma	4.00	2.00	[1.55, 12.4]	8.51	[7.13, 10.0]	21.9	[16.9, 27.9]
$\sigma_\mu$	$\mathbb{R}^+$	InvGamma	0.50	2.00	[0.20, 1.57]	2.22	[1.31, 3.07]	2.73	[1.76, 3.72]
$\sigma_g$	$\mathbb{R}^+$	InvGamma	0.75	2.00	[0.30, 2.32]	0.36	[0.33, 0.40]	0.58	[0.52, 0.63]
Marginal Likelihood						-1043.70		-1098.34	

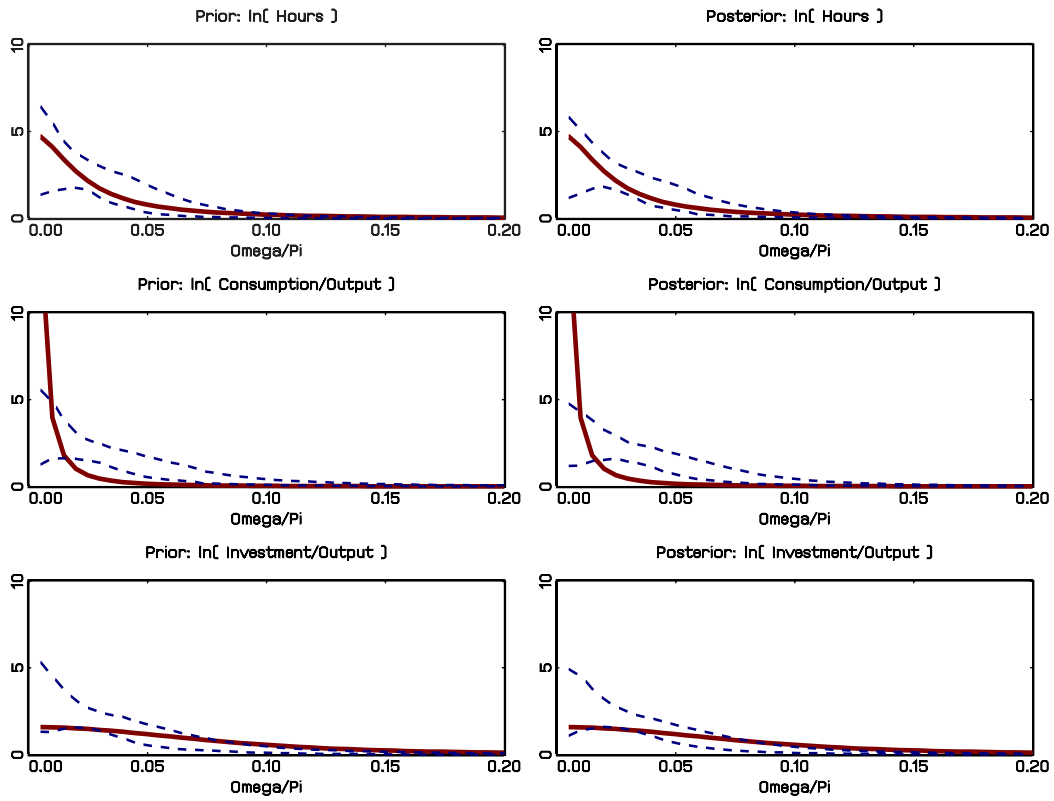
Notes:  $\mathcal{B}$  is Beta,  $\mathcal{G}$  is Gamma,  $\mathcal{IG}$  is Inverse Gamma, and  $\mathcal{N}$  is Normal distribution.  $P(1)$  and  $P(2)$  denote means and standard deviations for Beta, Gamma, and Normal distributions;  $s$  and  $\nu$  for the Inverse Gamma distribution, where  $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-\nu-1} e^{-\nu s^2/2\sigma^2}$ . The effective prior is truncated at the boundary of the determinacy region and the prior probability interval reflects this truncation. All probability intervals are 90% credible. The following parameters are fixed:  $\delta = 0.025$  and  $\beta = 1/(1 + 0.005)$ . Estimation results are based on the sample period QIV:1955 - QIV:2005.

Table 2: EXAMPLE 2: LOG MARGINAL DATA DENSITIES

$\ln p(Y)$		
$\lambda$	MCMC Approx	Exact
$\zeta = 1/4$		
1/10	-356.39	N/A
1	-356.63	-356.58
10	-360.06	N/A
$\zeta = 1/2$		
1/10	-353.24	N/A
1	-353.89	-353.90
10	-357.28	N/A
$\zeta = 3/4$		
1/10	-353.23	N/A
1	-355.58	-355.56
10	-357.51	N/A

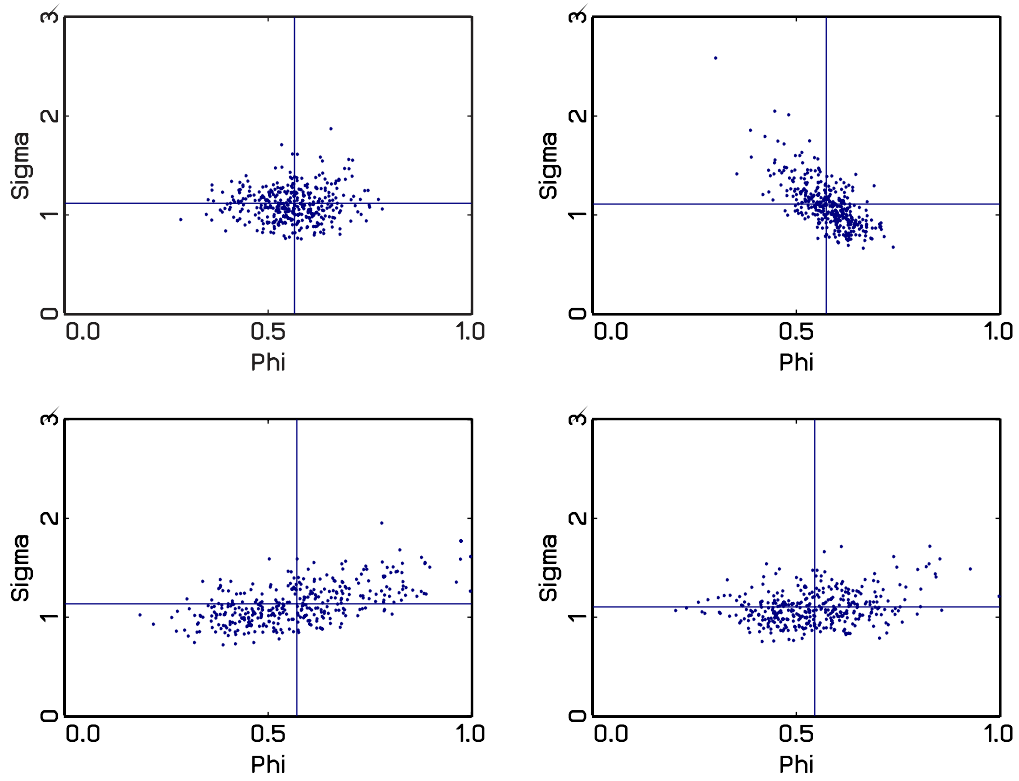
*Notes:* Results are based on a VAR(4), estimated with  $T = 120$  model generated data. For the MCMC Approx the prior density is set to zero for values of  $\Phi$  that imply non-stationarity.

Figure 1: THE “GREAT RATIOS” AND HOURS WORKED: PREDICTIVE DISTRIBUTIONS



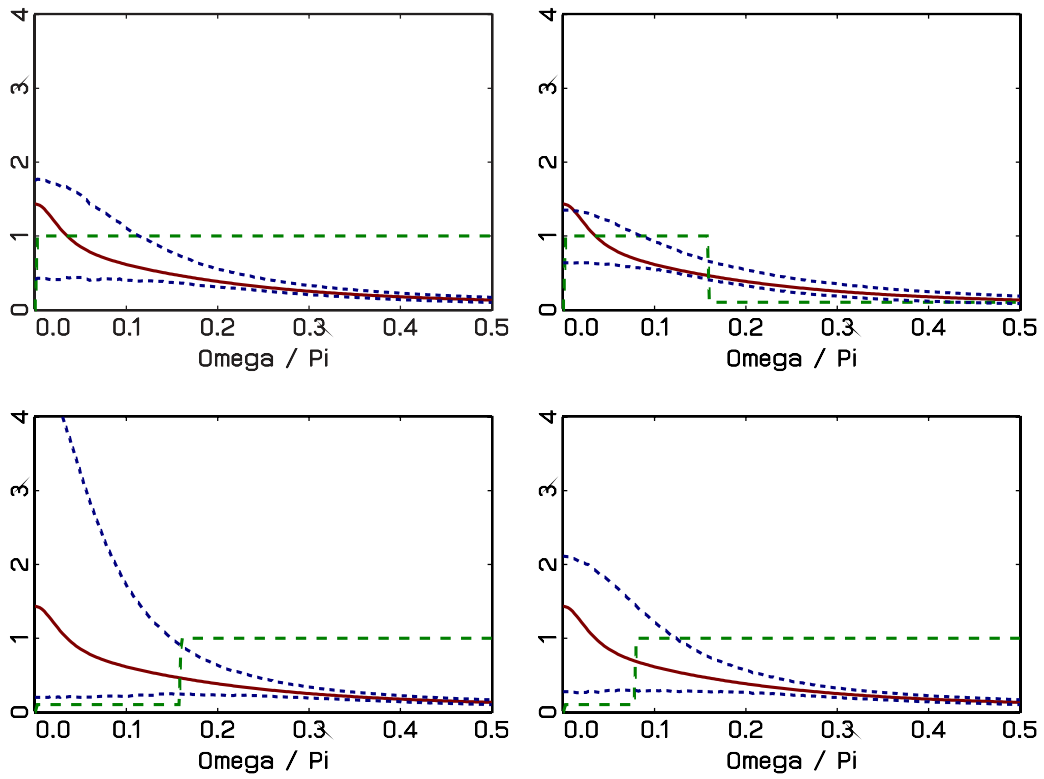
*Notes:* Figure depicts smoothed periodograms for the three normalized time series over the interval  $\omega/\pi \in [0.005, 0.200]$ : solid lines correspond to the actual data and dashed lines signify 90% probability bands from the prior and posterior predictive distributions under the DSGE model presented in Section 2.

Figure 2: EXAMPLE 1: PARAMETER DRAWS (EXACT)



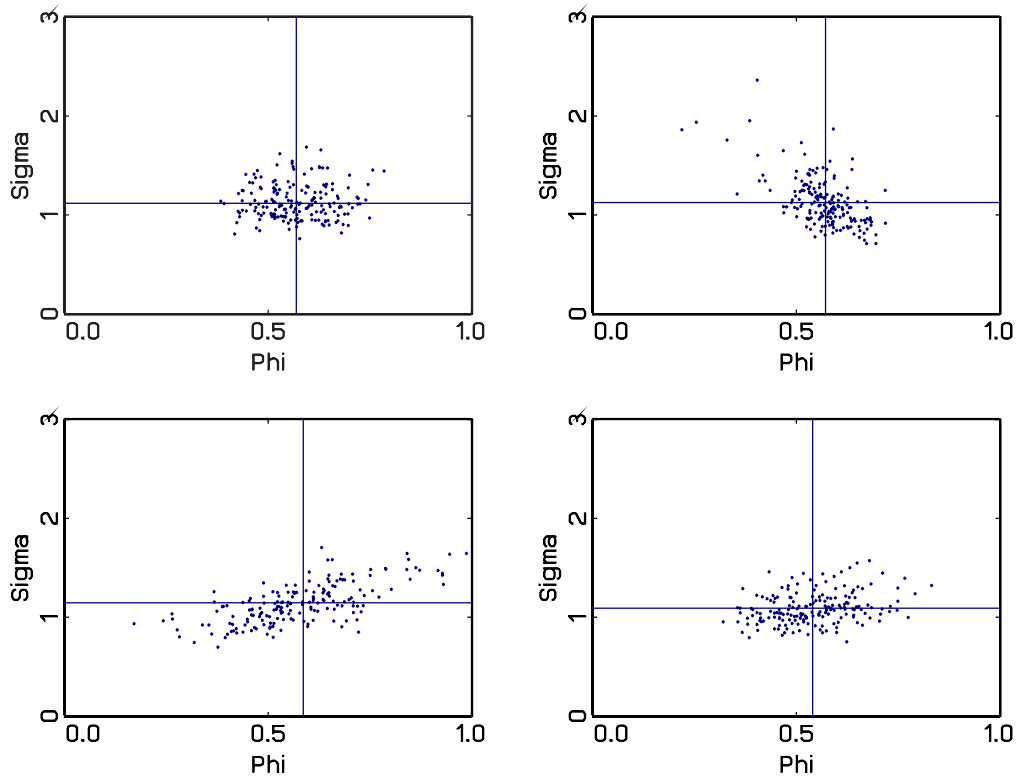
*Notes:* Figure depicts 200 draws from prior distribution for 4 different choices of  $\lambda(\omega)$ . Intersection of solid lines indicates prior mean. Panel (1,1) corresponds to a uniform  $\lambda(\omega)$ , in Panel (1,2) we emphasize frequencies below  $0.16\pi$ , in Panel (2,1) we emphasize frequencies above  $0.16\pi$ , and in Panel (2,2) we emphasize frequencies above  $0.08\pi$ .

Figure 3: EXAMPLE 1: SPECTRAL DENSITY DRAWS (EXACT)



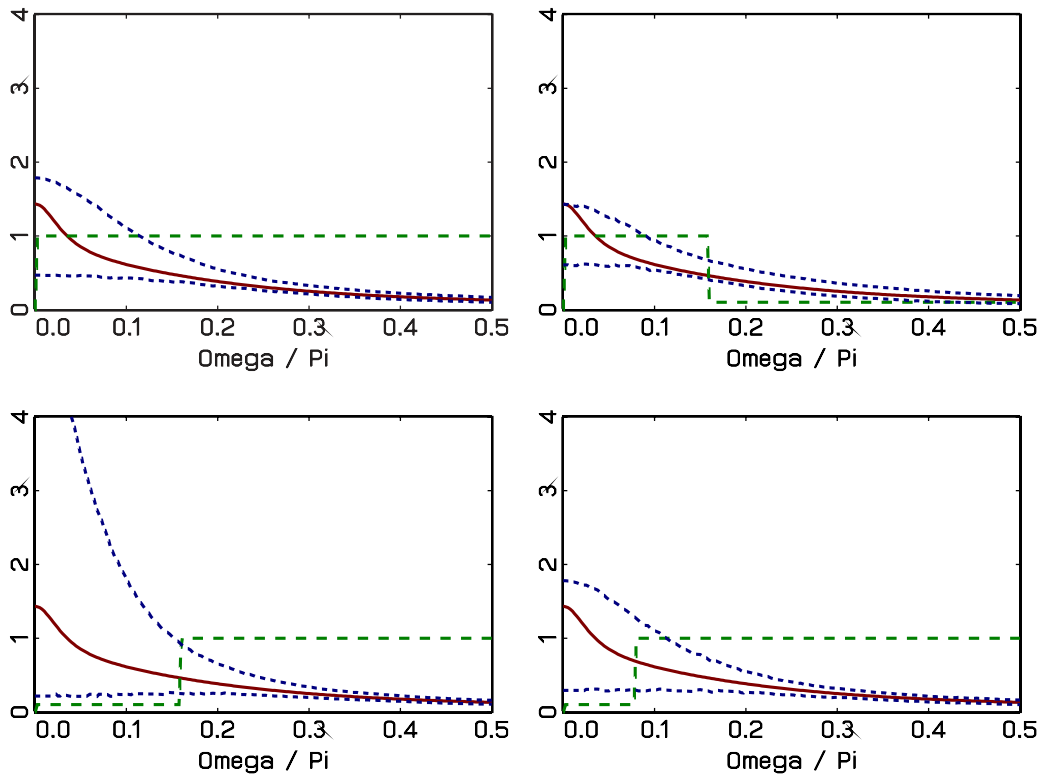
Notes: Figure depicts pointwise 90% probability intervals based on draws from the prior distribution of the spectral densities (short dashes) for 4 different choices of  $\lambda(\omega)$  (long dashes). The solid line indicates the target density  $S_D(\omega)$ .

Figure 4: EXAMPLE 1: PARAMETER DRAWS (APPROX)



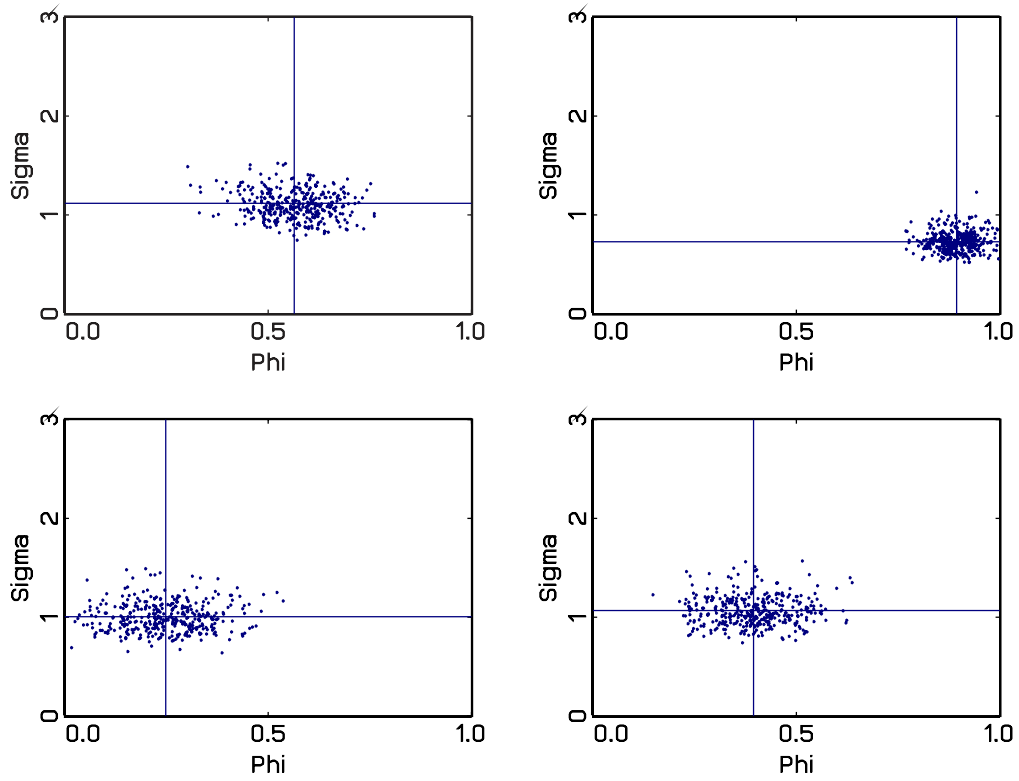
*Notes:* Figure depicts 200 draws from prior distribution for 4 different choices of  $\lambda(\omega)$ . Intersection of solid lines indicates prior mean. Panel (1,1) corresponds to a uniform  $\lambda(\omega)$ , in Panel (1,2) we emphasize frequencies below  $0.16\pi$ , in Panel (2,1) we emphasize frequencies above  $0.16\pi$ , and in Panel (2,2) we emphasize frequencies above  $0.08\pi$ .

Figure 5: EXAMPLE 1: SPECTRAL DENSITY DRAWS (APPROX)



Notes: Figure depicts pointwise 90% probability intervals based on draws from the prior distribution of the spectral densities (short dashes) for 4 different choices of  $\lambda(\omega)$  (long dashes). The solid line indicates the target density  $S_D(\omega)$ .

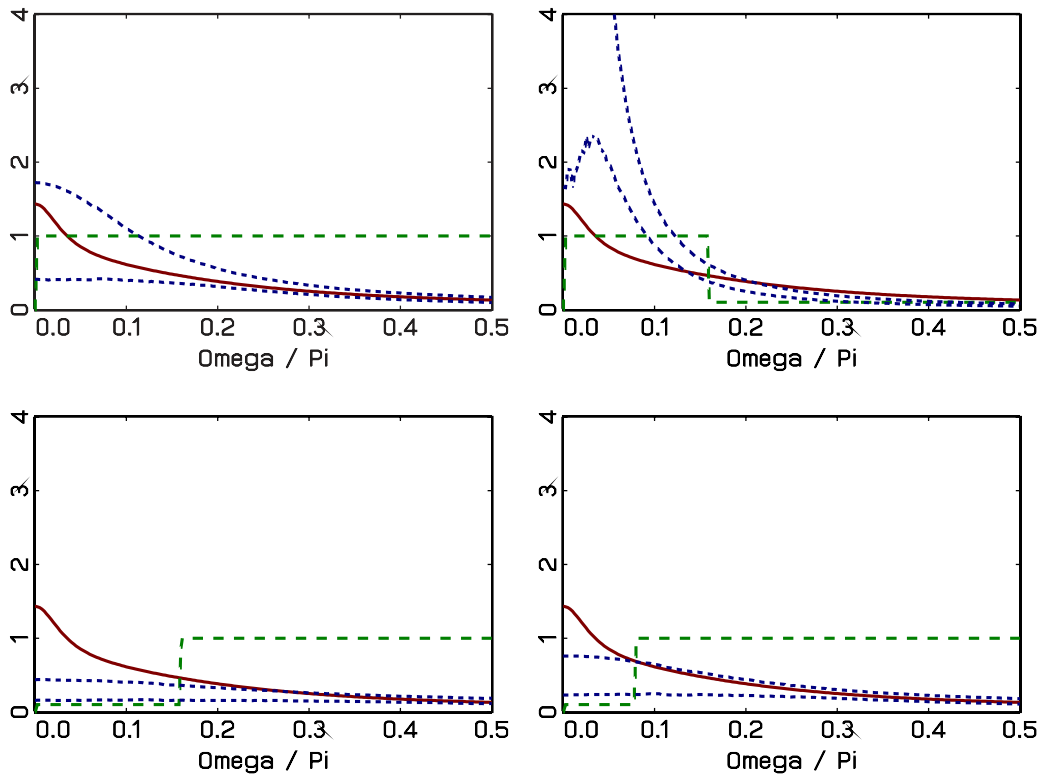
Figure 6: EXAMPLE 1: PARAMETER DRAWS (BANDPASS-FILTERED DUMMIES)



*Notes:* Figure depicts 200 draws from prior distribution for 4 different choices of  $\lambda(\omega)$ . Intersection of solid lines indicates prior mean. Panel (1,1) corresponds to a uniform  $\lambda(\omega)$ , in Panel (1,2) we emphasize frequencies below  $0.16\pi$ , in Panel (2,1) we emphasize frequencies above  $0.16\pi$ , and in Panel (2,2) we emphasize frequencies above  $0.08\pi$ .



Figure 7: EXAMPLE 1: SPECTRAL DENSITY DRAWS (BANDPASS-FILTERED DUMMIES)



Notes: Figure depicts pointwise 90% probability intervals based on draws from the prior distribution of the spectral densities (short dashes) for 4 different choices of  $\lambda(\omega)$  (long dashes). The solid line indicates the target density  $S_D(\omega)$ .

Figure 8: EXAMPLE 2: DSGE AND DGP SPECTRAL DENSITIES

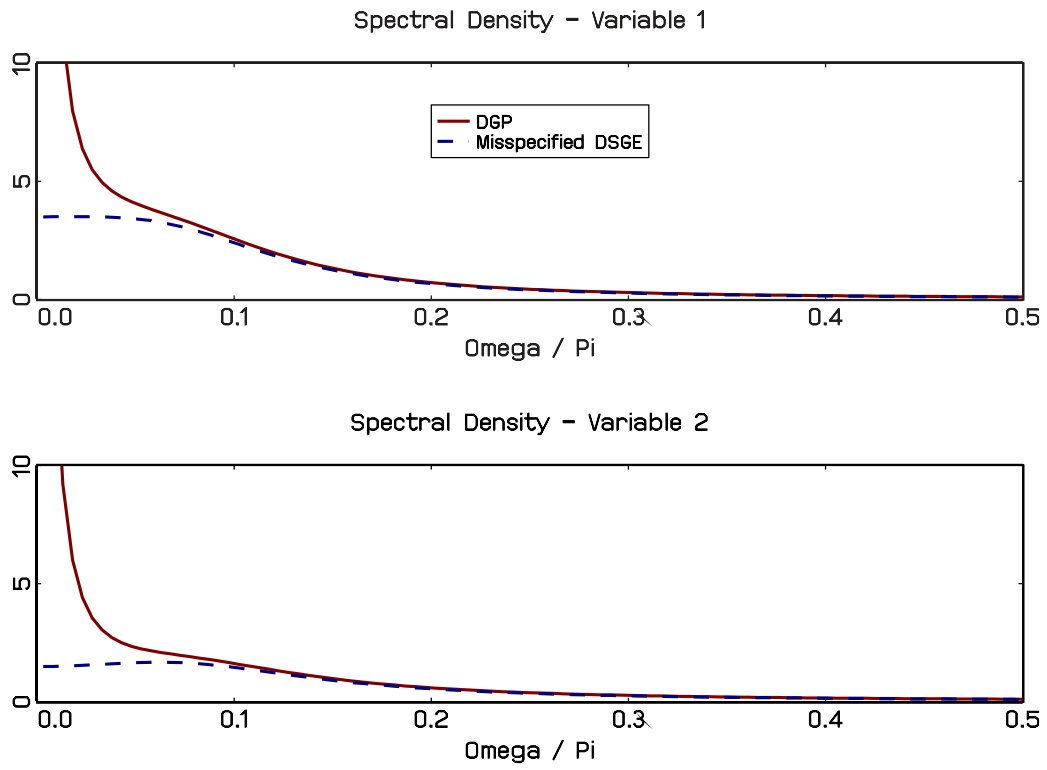
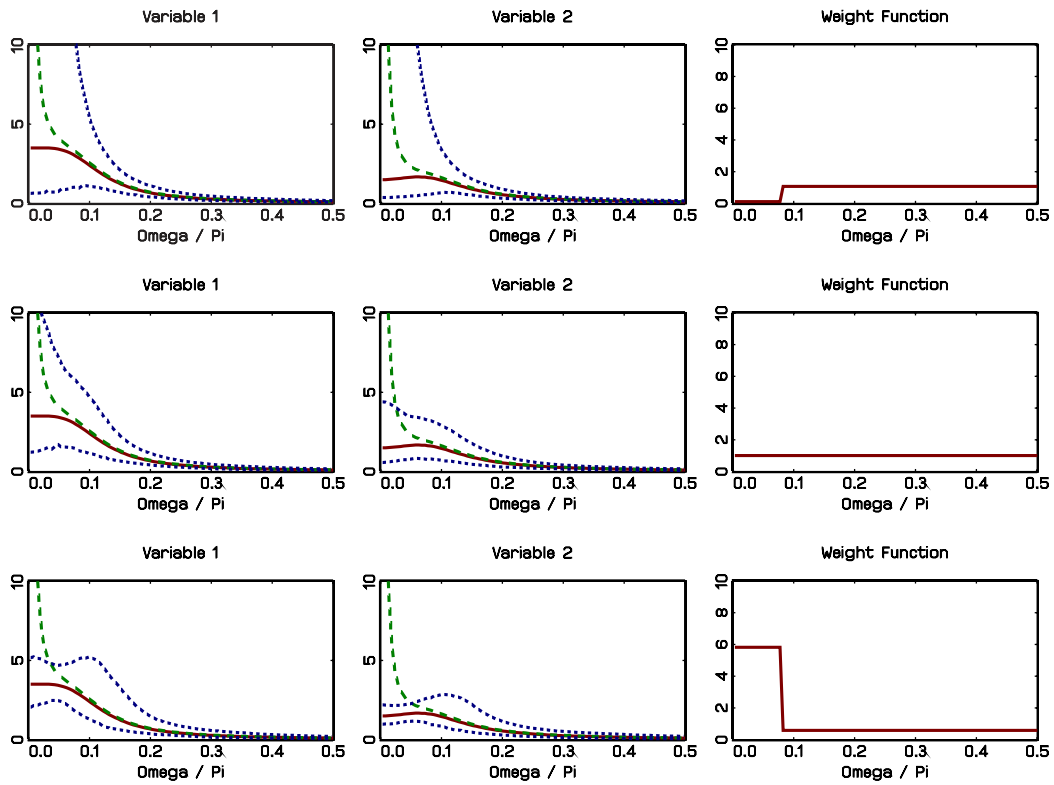
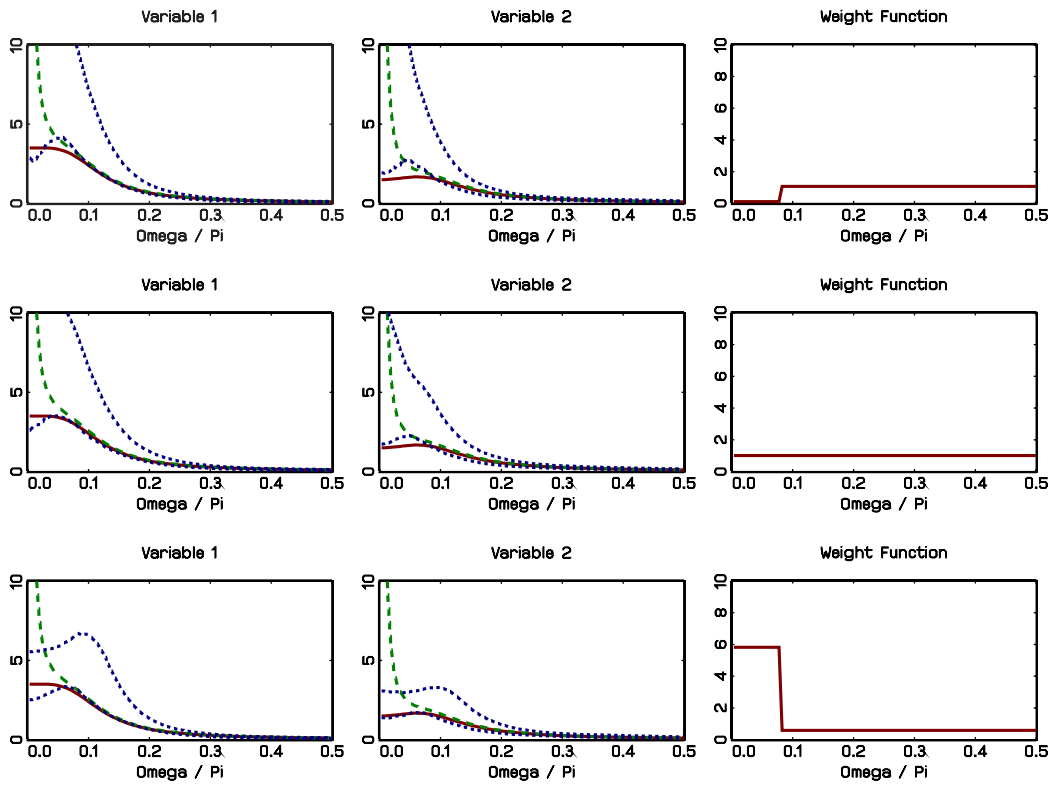


Figure 9: EXAMPLE 2: PRIOR DISTRIBUTION OF SPECTRUM



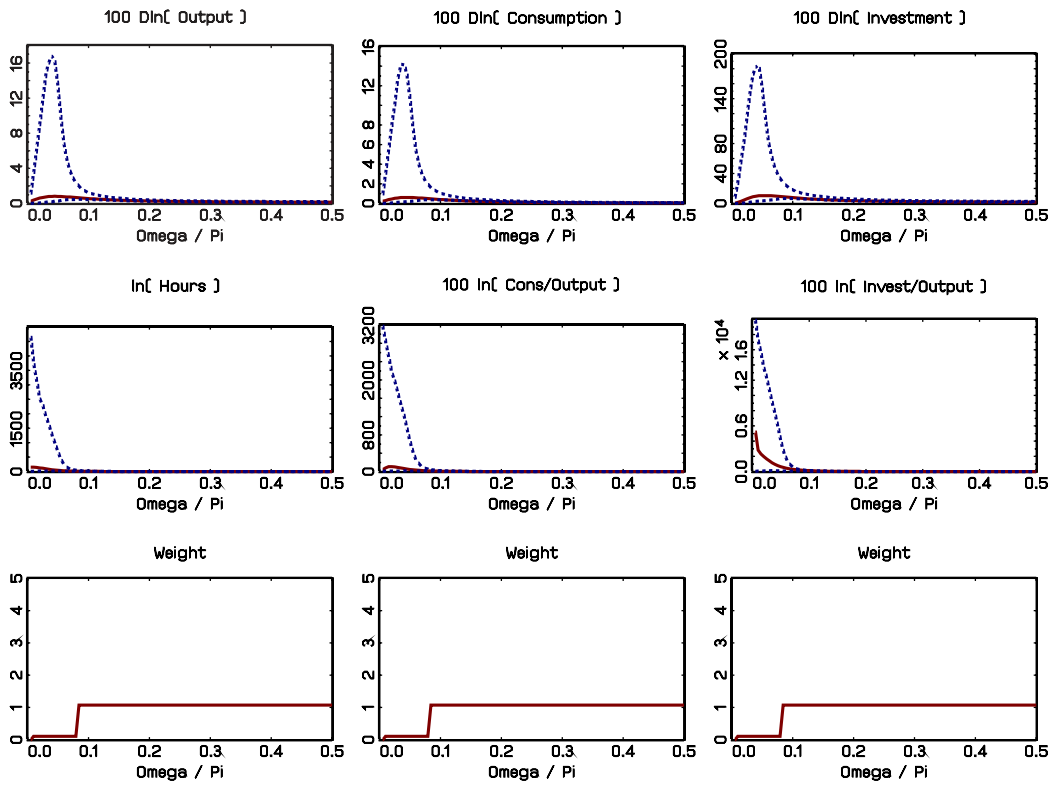
Notes: Figure depicts pointwise 90% probability intervals based on draws from the prior distribution of the spectral densities (short dashes) for 3 different choices of  $\lambda(\omega)$  (right column). The solid line indicates the target spectrum  $S_D(\omega)$  and the long dashes show the spectrum of the DGP.

Figure 10: EXAMPLE 2: POSTERIOR DISTRIBUTION OF SPECTRUM



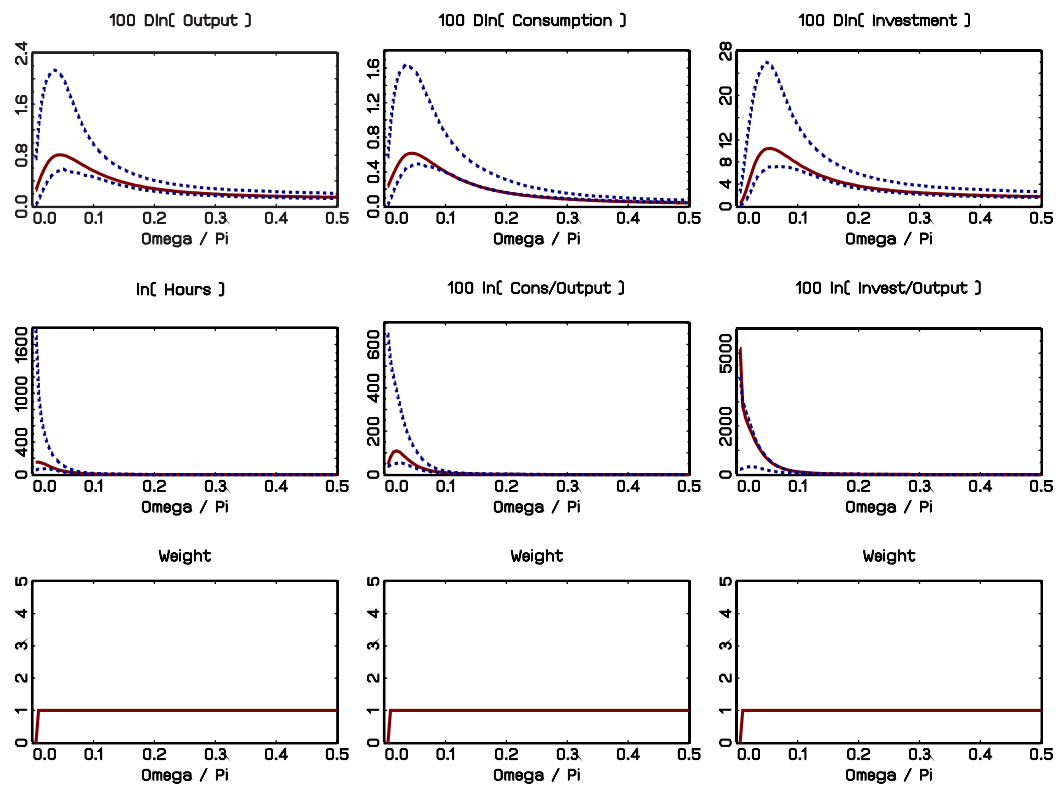
Notes: Figure depicts pointwise 90% probability intervals based on draws from the prior distribution of the spectral densities (short dashes) for 3 different choices of  $\lambda(\omega)$  (right column). The solid line indicates the target spectrum  $S_D(\omega)$  and the long dashes show the spectrum of the DGP.

Figure 11: DSGE-VAR: PRIOR FOR SPECTRUM, EMPHASIZE BUSINESS CYCLE



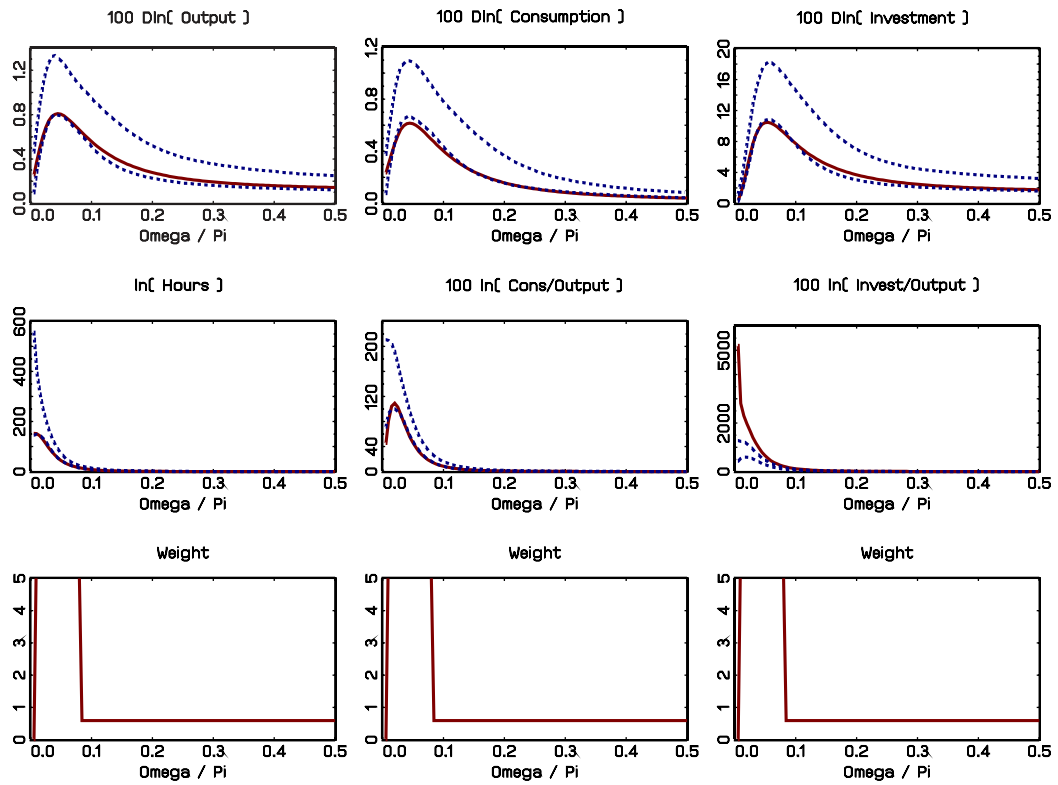
Notes: Figure depicts pointwise 90% probability intervals of the prior predictive distribution (short dashes). The solid line indicates the sample spectrum.

Figure 12: DSGE-VAR: PRIOR FOR SPECTRUM, EQUAL WEIGHTS



Notes: Figure depicts pointwise 90% probability intervals of the prior predictive distribution (short dashes). The solid line indicates the sample spectrum.

Figure 13: DSGE-VAR: PRIOR FOR SPECTRUM, EMPHASIZE LONG-RUN



Notes: Figure depicts pointwise 90% probability intervals of the prior predictive distribution (short dashes). The solid line indicates the sample spectrum.