Labor-Supply Shifts and Economic Fluctuations

Technical Appendix*

Yongsung Chang
Department of Economics
University of Pennsylvania†

Frank Schorfheide
Department of Economics
University of Pennsylvania‡

January 2003

---

*Not intended for publication. The GAUSS programs to implement the empirical analysis are available at *http://www.econ.upenn.edu/~schorf*.

†McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104-6297. Tel.: (215) 898-6691. Email: *yohg@ssc.upenn.edu*.

‡Tel.: (215) 898-8486. Email: *schorf@ssc.upenn.edu*.

1 VAR Specification

Define the vector of stationary variables $\Delta y_t = [\Delta \ln P_t, \Delta \ln I_{h,t}, \ln L_{m,t}]'$. Moreover, let $\epsilon_t = [\epsilon_{z,t}, \epsilon_{a,t}, \epsilon_{b,t}]'$. The VAR can be expressed in vector error correction form as

$$\Delta y_t = \Phi_0 + \Phi_{vec} y_{t-1} + \sum_{i=1}^{p-1} \Phi_i \Delta y_{t-i} + u_t, \quad u_t \sim iid\ \mathcal{N}(0, \Sigma_u). \tag{1}$$

The reduced form disturbances u_t are related to the structural disturbances ϵ_t by $u_t = \Phi_* \tilde{\epsilon}_t$, where $\tilde{\epsilon}_t$ is a standardized version of ϵ_t with unit variance.

1.1 Long-run Restrictions

Let $\Phi_{vec} = \mu\lambda'$, where μ is a 3×1 vector and $\lambda' = [1, -\lambda_{21}, 0]$. The VAR specification ensures that productivity and home investment have a common stochastic trend and the cumulative hours process $\tilde{L}_{mt}$ has a second stochastic trend. Let $\mu_\perp$ be a 3×2 matrix with columns that are orthogonal to μ and define $\lambda_\perp$ as matrix with columns $[\lambda_{12}, 1, 0]'$ and $[0, 0, 1]'$.

The stochastic trend of y_t has the form $C_{LR}\Phi_* \sum_{\tau=0}^{t} \tilde{\epsilon}_t$. The long-run multiplier matrix is given by

$$C_{LR} = \lambda_\perp \left[\mu_\perp' \left(I_{3\times 3} - \sum_{i=1}^{p} \Phi_i \right) \lambda_\perp \right]^{-1} \mu_\perp', \tag{2}$$

see for instance Theorem 4.2 in Johansen (1995). $I_{3\times 3}$ denotes the 3×3 identity matrix. Since productivity and consumption expenditures have a common trend, the first two rows of the 3×3 long-run multiplier matrix C_{LR} are proportional. The factor of proportionality is λ_{12}.

1.2 Bayesian Inference with Non-Identifiable Parameter

In the context of the VAR the structural shocks can only be determined if the slope $\alpha - 1$ of the inverse-labor-demand schedule is given. However, the likelihood function is uninformative with respect to α as it only depends on the reduced-form parameters $\theta_{(0)}$, that is,

$$p(Y_T|\theta_{(0)}, \alpha, \mathcal{M}_0) = \tilde{p}(Y_T|\theta_{(0)}, \mathcal{M}_0). \tag{3}$$

The joint posterior density of $\theta_{(0)}$ and α can be expressed as

$$\begin{aligned}
p(\theta_{(0)}, \alpha|Y_T, \mathcal{M}_0) &= \frac{\tilde{p}(Y_T|\theta_{(0)}, \mathcal{M}_0)p(\theta_{(0)}|\mathcal{M}_0)p(\alpha)}{\int \left[\tilde{p}(Y_T|\theta_{(0)}, \mathcal{M}_0)p(\theta_{(0)}|\mathcal{M}_0)[\int p(\alpha)d\alpha]\right] d\theta_{(0)}} \\
&= \frac{\tilde{p}(Y_T|\theta_{(0)}, \mathcal{M}_0)p(\theta_{(0)}|\mathcal{M}_0)}{\int \left[\tilde{p}(Y_T|\theta_{(0)}, \mathcal{M}_0)p(\theta_{(0)}|\mathcal{M}_0)\right] d\theta_{(0)}} p(\alpha) \\
&= \tilde{p}(\theta_{(0)}|Y_T, \mathcal{M}_0)p(\alpha). \qquad (4)
\end{aligned}$$

Thus, $p(\theta_{(0)}, \alpha|Y_T, \mathcal{M}_0)$ is the product of the posterior density of the (identifiable) reduced form parameters and the prior density of α. According to the VAR the data Y_T convey no information about α. Hence, the prior density $p(\alpha)$ is not updated after observing Y_T.[1]

1.3 Prior

Let ΔY_T be the $(T-p) \times n$ matrix with rows $\Delta y_t'$, $t = p+1, \ldots, T$ (the first p observations are used to initialize lags). Let $k = 3 + np$, $X_T(\lambda_{21})$ be the $(T-p) \times k$ matrix with rows $x_t' = [1, t, (1, -\lambda_{21}, 0)y_{t-1}, \Delta y_{t-1}', \ldots, \Delta y_{t-p}']$, U_T be the matrix with rows u_t', and $B = [\Phi_0, \Phi_{tr}, \alpha, \Phi_1, \ldots, \Phi_p]'$. We include a deterministic trend with coefficient vector Φ_{tr} in the specification of $\mathcal{M}_0$ to capture potential long-run shifts in market hours due to structural changes in labor market participation behavior. The VAR can be expressed in matrix form as $\Delta Y_T = X_T(\lambda_{21})B + U_T$. Conditional on λ_{21} the prior for B and Σ is constructed from a training sample

[1] Suppose $\mathcal{M}_0$ and $\mathcal{M}_1$ are analyzed jointly by placing prior probabilities $\pi_{i,0}$ on the two models. Despite the presence of the DSGE model $\mathcal{M}_1$ and the informative posterior $p(\alpha|Y_T, \mathcal{M}_1)$ that it generates, it is still true that the VAR impulse responses have to be identified through the prior $p(\alpha)$, not the DSGE model posterior $p(\alpha|Y_T, \mathcal{M}_1)$, or the overall marginal posterior $p(\alpha|Y_T) = \pi_{0,T}p(\alpha) + \pi_{1,T}p(\alpha|Y_T, \mathcal{M}_1)$.

$t = p+1, \ldots, T_*$. Let ΔY_* and $X_*(\lambda_{21})$ be matrices with rows $\Delta y_t'$ and x_t' as defined above, $t = p+1, \ldots, T_*$. Define

$$\hat{B}_* = (X_*'X_*)^{-1}X_*'\Delta Y_*, \quad \hat{\Sigma}_{u,*} = (T_* - p)^{-1}(Y_* - X_*\hat{B}_*)'(Y_* - X_*\hat{B}_*). \tag{5}$$

Then we obtain

$$\begin{aligned}
\lambda_{21} &\sim \mathcal{N}(1, 0.025^2) \\
\Sigma_u|\lambda_{21}, Y_* &\sim IW\Big((T_* - p)\hat{\Sigma}_{u,*}, T_* - k - p\Big) \\
vec(B)|\Sigma_u, \lambda_{21}, Y_* &\sim \mathcal{N}\Big(vec(\hat{C}_*), \Sigma_u \otimes (X_*'X_*)^{-1}\Big),
\end{aligned} \tag{6}$$

where IW denotes the Inverted Wishart distribution. In our empirical analysis the size of the training sample is $T_* = 20$ and the lag-length is $p = 2$.

1.4 Posterior Simulation

A Gibbs sampler is used to generate 110,000 draws from the posterior distribution of the VAR parameters $(B, \Sigma_u, \lambda_{21})$. We draw successively from the conditional posteriors $p(B, \Sigma_u|\lambda_{21}, Y_T, \mathcal{M}_0)$ and $p(\lambda_{21}|B, \Sigma_u, Y_T, \mathcal{M}_0)$. The distribution of $\Sigma_u|\lambda_{21}, Y_T$ is Inverted Wishart and $B|\Sigma_u, \lambda_{21}, Y_T$ is multivariate normal. The parametrization is given by replacing ΔY_* and $X_*(\lambda_{21})$ with ΔY_T and X_T in Equations (5) and (6). To characterize the posterior distribution of λ_{21}, define $\Delta\tilde{Y}_T$ with rows $\Delta\tilde{y}_t' = [\Delta y_t - \Phi_0 - \Phi_{tr}t - \mu(1,0,0)y_{t-1} - \sum_{i=1}^p \Phi_i \Delta y_{t-i}]'$ and $\tilde{X}_t$ with rows $\tilde{x}_t'$, where $\tilde{x}_t = \mu(0,-1,0)y_{t-1}$. Then one obtains

$$\lambda_{21}|B, \Sigma_u, Y_T \sim \mathcal{N}(m_\lambda, v_\lambda), \tag{7}$$

where $v_\lambda^{-1} = 1/0.01 + tr[\Sigma^{-1}\tilde{X}_T'\tilde{X}_T]$, $m_\lambda = v_\lambda\left(\frac{1}{0.01} + \frac{tr[\Sigma^{-1}\Delta\tilde{Y}_T'\tilde{X}_T]}{tr[\Sigma^{-1}\Delta\tilde{X}_T'\tilde{X}_T]}\right)$, and $tr[.]$ denotes the trace operator. The first 10,000 draws are discarded. The parameter draws are converted into population values of the (integrated) spectral densities and truncated impulse response functions. The marginal data density of the DSGE model is approximated with Geweke's (1999) modified harmonic mean estimator.

2 DSGE Model Analysis

The log-linearized DSGE model is solved by standard methods. Conditional on parameter values $\theta_{(1)}, \alpha$, the likelihood function of the log-linearized DSGE model can be evaluated with the Kalman filter. A numerical-optimization routine is used to find the posterior mode. The inverse Hessian is calculated at the posterior mode. 110,000 draws from the posterior distribution of the DSGE model parameters are generated with a random-walk Metropolis-Hastings algorithm. The scaled inverse Hessian serves as a covariance matrix for the Gaussian proposal distribution used in the Metropolis-Hastings sampler. The first 10,000 draws are discarded. The parameter draws are converted into population values of the (integrated) spectral densities and truncated impulse response functions. The marginal data density of the DSGE model is approximated with Geweke's (1999) modified harmonic mean estimator. Details of these computations are discussed in the appendix of Schorfheide (2000).

For each draw from the posterior distribution of DSGE model parameters $[\theta'_{(1)}, \alpha]'$ a smoothing algorithm is applied to compute expected values for the technology sequences $\{a_t\}_{t=1}^T$, $\{b_t\}_{t=1}^T$, and $\{z_t\}_{t=1}^T$ conditional on $\mathcal{M}_1$ and the sample of observations Y_T. These sequences of expected values are averaged across the parameter draws and plotted in Figure 2 of the paper.

3 A Small Simulation Experiment

The VAR identification scheme is "consistent" with the DSGE model in the following sense. (We only give a heuristic argument) Suppose data from the DSGE model are generated as follows: (1) generate a draw of parameters $\theta_{(1)0}, \alpha_0$ from the prior distribution. (2) Generate a sample of size T from the DSGE model, using the parameters generated in Step (1). Then the VAR is fitted and the identification procedure is applied. The DSGE model has an infinite-order moving average (MA) representation in terms of the structural shocks. Up to the relationship between the structural shocks and the reduced form VAR innovations, $u_t = \Phi_* \tilde{\epsilon}_t$, this MA representation can be estimated with more and more precision as the sample size T increases, provided that the lag-length of the VAR is increased appropriately as the size of the sample grows.[2]

If α_0 were known, then one could asymptotically recover the true MA representation of the DSGE model because the VAR identification is based on restrictions that are satisfied by the DSGE model. (See discussion in the paper). However, α_0 is not known and therefore the posterior distribution of the MA representation does not degenerate to a point mass asymptotically. However, since α_0 must lie in the support of the prior for α, the "true" MA representation must lie in (or at least very close to) the support of the posterior distribution of the MA representation obtained from our VAR analysis. Our posterior will assign high probabilities to MA representations that correspond to α's with high prior probability.

To illustrate the VAR identification procedure and the effect of the non-identifiability of the parameter α, a small simulation experiment is conducted. Data is generated from the DSGE model $\mathcal{M}_1$. Posterior mean estimates obtained in the empirical analysis are used to parameterize the DSGE model. In particular, α is set equal to 0.74. We use sample sizes $T = 20$ and $T = 5000$. The former corresponds to the length of the pre-sample that is used in the empirical analysis to set the prior for the reduced form VAR parameters. The latter sample size, much larger than the typical macroeconomic data set, will highlight the large sample characteristic of our approach. Based on the artificial data we estimate VARs and generate a posterior distributions for the variance decomposition of output at the frequency 1/12 cycles per quarter.

As in the actual empirical analysis, the prior mean of α is chosen to be 0.66. Two different values for the prior standard deviation of α are used. The value $\sigma(\alpha) = 0.02$ implies a 95 percent confidence interval ranging from 0.62 to 0.70. This interval is consistent with a short sample of postwar U.S. labor income shares. The value $\sigma(\alpha) = 0.2$ leads to a confidence interval from 0.46 to 1.06, which covers most plausible as well as many implausible values of α.

[2]Most DSGE models cannot be written as finite-order VARs.

Figure 1 visualizes the variance decomposition of output at frequency 1/12 obtained from simulated data. Since the variance decompositions have to sum to one across shocks, they lie in a two dimensional triangular shaped subspace (simplex) of $I\!R^3$. Each dot in the four panels of Figures 1 corresponds to a draw from the posterior distribution of the variance decomposition based on the VAR. A dot d can be converted in numerical value for the relative contribution of a labor supply shock to the variance of output as follows:

$$\text{contribution of b-shock} = \frac{\text{Area of triangle z-d-a}}{\text{Area of triangle z-b-a}}.$$

Hence, the three corner b of the simplexes corresponds to the decomposition that assigns 100 percent of the variation to the labor supply shock, and 0 percent to the other two shocks. Clusters of points indicate regions of high posterior density. V signifies the posterior mean for the VAR.

Informal inspection of the plots suggests that for small samples, such as $T = 20$, the uncertainty with respect to the variance decomposition is dominated by the uncertainty about the reduced form VAR parameters. The dispersion of the posterior draws is quite similar for both choices of $\sigma(\alpha)$. As the sample size is increased to $T = 5000$, the posterior variance of the identifiable VAR parameters decreases substantially. Nevertheless, there remains substantial uncertainty about the role of permanent versus temporary market technology shocks. A drawback of the long-run identification restriction is that it leads to imprecise decompositions.

This paper focuses on the role of the labor demand shock versus the two technology shocks. For $\sigma(\alpha) = 0.02$ the posterior uncertainty, reflected by the vertical spread of the draws, is very small. If $\sigma(\alpha)$ is increased to 0.2 the spread becomes larger. Nevertheless, a comparison of Panel 1 with 3, and Panel 2 with 4 shows that the sample information leads to an update of the beliefs about the relative importance of labor supply shocks. Due to the non-identifiability of α, the proposed identification procedure is not consistent in the sense that the posterior degenerates to the "true" decomposition that corresponds to the parameterized DSGE model as the sample size approaches infinity. Nevertheless, it enables the researcher to extract information from the data and learn about impulse responses and variance decompositions. Under a tight prior for α, e.g. $\sigma(\alpha) = 0.02$, which we think is justified in the analysis of the paper, our procedure will lead to a concentrated posterior in a large sample.

References

[1] Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin (1995) *"Bayesian Data Analysis,"* Chapman & Hall, New York.

[2] Geweke, J. (1999) "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," Econometric Reviews, 18, 1-126.

[3] Johansen, S. (1995) "Likelihood-Based Inference in Cointegrated Vector Autoregressive Models," Oxford University Press, Oxford.

[4] Schorfheide, F. (2000) "Loss Function Based Evaluation of DSGE Models," Journal of Applied Econometrics, 15, 645-670.

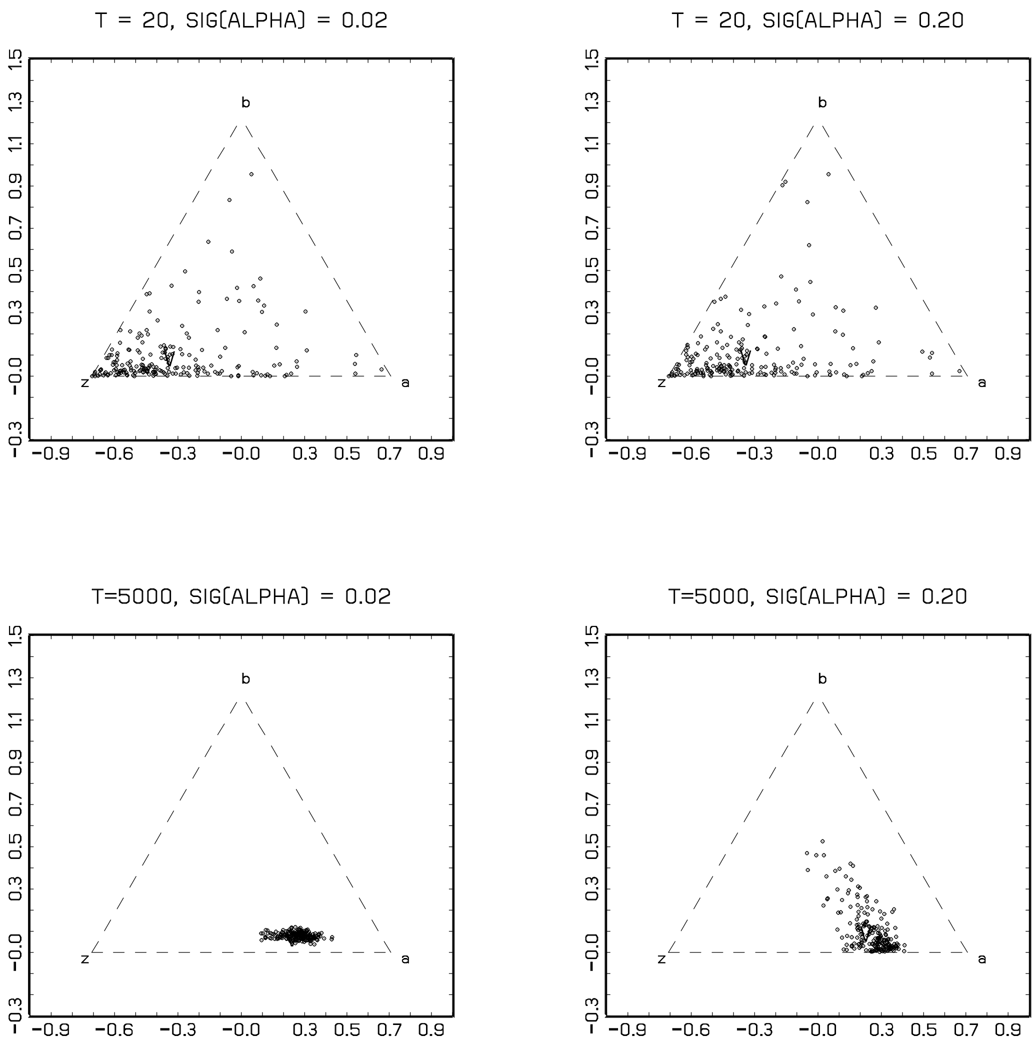

Figure 1: Spectral decomposition of output at frequency 1/12 cycles per quarter based on artificial observations generated from DSGE model. Sample size: $T = 20$ and $T = 5000$. Prior standard errors of α are 0.02 and 0.20, respectively. Dots correspond to 200 draws from VAR posterior distribution. V indicates posterior mean of VAR.