Bayesian Analysis of DSGE Models

Frank Schorfheide
Department of Economics, University of Pennsylvania

Comparison to VARs: DSGE-VARs

- Compare fit of DSGE model to that of a VAR based on marginal data densities. Mechanics are non-trivial. Under a very diffuse prior for the VAR coefficients, the DSGE model is likely to win the comparison.
- Careful construction of VAR prior is crucial, for instance:
  - Minnesota-style prior, Sims-Zha priors for identified VARs.
  - DSGE-VARs: Del Negro and Schorfheide (2004, 2005), Del Negro, Schorfheide, Smets, and Wouters (2004).
- Compare DSGE model dynamics to (identified) VAR dynamics

Two Views of DSGE-VARs

- Improve VAR estimates by “restricting” its parameter estimates.

- Improve DSGE model by relaxing its restrictions.

DSGE-VARs: Improving VARs

- Consider a vector autoregressive specification of the form

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + u_t, \quad I\!E[u_t u_t'] = \Sigma. \tag{1}$$

- Write VAR as $Y = X\Phi + U$, Y is $T \times n$, X is $T \times k$.

- Difficulty: too many parameters which leads to imprecise estimates.

- Solution: tilt estimates toward a point in the parameter space. Example: Minnesota prior tilts toward random walks.

- Here: tilt toward DSGE model restrictions.

Example

- n independent draws of y_t from $\mathcal{N}(\mu, 1)$.

- MLE of μ:

$$\hat{\mu}_{ML} = \frac{1}{n} \sum_{t=1}^{n} y_t.$$

- Bayes estimator based on prior $\mu \sim \mathcal{N}(0, \tau^2)$

$$\hat{\mu}_B = \frac{1}{n + 1/\tau^2} \sum_{t=1}^{n} y_t = \frac{n}{n + 1/\tau^2} \hat{\mu}_{ML} + \frac{1/\tau^2}{n + 1/\tau^2} 0$$

Example

- MSE of MLE:

$$\mathbb{E}_\mu\left[(\mu - \hat{\mu}_{ML})^2\right] = \underbrace{0^2}_{\text{Bias}^2} + \underbrace{\frac{1}{n}}_{\text{Variance}} .$$

- MSE of Bayes Estimator:

$$\mathbb{E}_\mu\left[(\mu - \hat{\mu}_B)^2\right] = \underbrace{\mu^2 \left(\frac{1/\tau^2}{n + 1/\tau^2}\right)^2}_{\text{Bias}^2} + \underbrace{\frac{n}{(n + 1/\tau^2)^2}}_{\text{variance}} .$$

- If μ^2 is small, i.e. the discrepancy between the *a priori* expected value and the "true" value is small, then the Bayes estimator clearly dominates.

DSGE-VARs: Improving VARs

- Complication: DSGE model depends on parameters θ.
- Solution: place prior on θ. Use notion of dummy observations to construct priors conditional on θ. Overall:

$$p(\Phi, \Sigma, \theta) = p(\theta)p(\Phi, \Sigma|\theta). \tag{2}$$

- Let's look at: $p(\Phi, \Sigma|\theta)$.

DSGE-VARs: Improving VARs

- Quasi-likelihood function for artificial observations (sample size $T^* = \lambda T$) generated from DSGE model:

$$p(Y^*(\theta)|\Phi, \Sigma_u) \propto \tag{3}$$
$$|\Sigma_u|^{-\lambda T/2} \exp\left\{ -\frac{1}{2} tr[\Sigma_u^{-1}(Y^{*'}Y^* - \Phi' X^{*'} Y^* - Y^{*'} X^* \Phi + \Phi' X^{*'} X^* \Phi)] \right\}.$$

- Let $I\!E^D_\theta[\cdot]$ be the expectation under DSGE model and define the autocovariance matrices

$$\Gamma_{XX}(\theta) = I\!E^D_\theta[x_t x_t'], \quad \Gamma_{XY}(\theta) = I\!E^D_\theta[x_t y_t'].$$

- Replace sample moments $Y^{*'}Y^*$ by $I\!E^D_\theta[Y^{*'}Y^*] = \lambda T \Gamma_{YY}(\theta)$, etc.

DSGE-VARs: Improving VARs

- Define

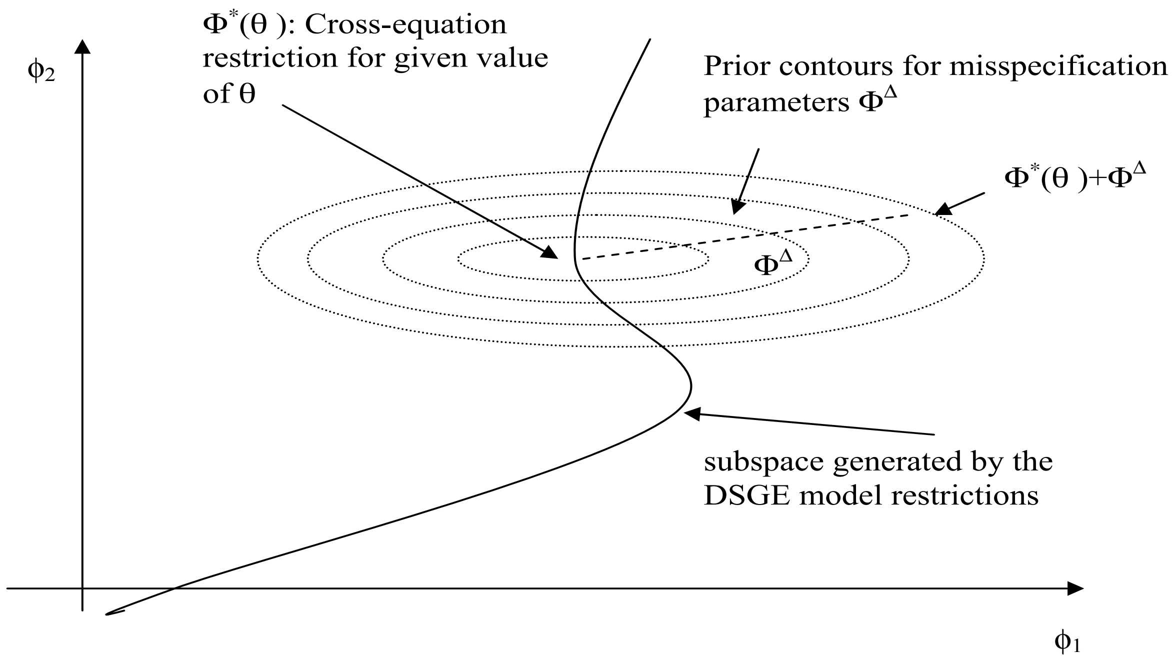
$$\Phi^*(\theta) = \Gamma^{-1}_{XX}(\theta)\Gamma_{XY}(\theta), \quad \Sigma^*(\theta) = \Gamma_{YY}(\theta) - \Gamma_{YX}(\theta)\Gamma^{-1}_{XX}(\theta)\Gamma_{XY}(\theta). \tag{4}$$

- Prior distribution:

$$\begin{aligned} \Sigma|\theta \;&\sim\; \mathcal{IW}\bigg(\lambda T\Sigma^*(\theta), \lambda T - k, n\bigg) \qquad (5)\\ \Phi|\Sigma,\theta \;&\sim\; \mathcal{N}\left(\Phi^*(\theta), \frac{1}{\lambda T}\bigg[\Sigma^{-1} \otimes \Gamma_{XX}(\theta)\bigg]^{-1}\right), \end{aligned}$$

DSGE-VARs: Relaxing DSGE Restrictions

- Alternative motivation...



$\Phi^*(\theta)$: Cross-equation restriction for given value of θ
ϕ_2
Prior contours for misspecification parameters Φ^Δ
$\Phi^*(\theta)+\Phi^\Delta$
Φ^Δ
subspace generated by the DSGE model restrictions
ϕ_1

DSGE-VARs: Relaxing DSGE Restrictions

- There is a vector θ and matrices Φ^Δ and Σ^Δ such that the data are generated from the VAR in Eq. (1)

$$\Phi = \Phi^*(\theta) + \Phi^\Delta, \quad \Sigma = \Sigma^*(\theta) + \Sigma^\Delta. \tag{6}$$

- We will construct a prior for Φ^Δ and Σ^Δ
- For now assume $\Sigma^\Delta = 0$.

DSGE-VARs: Relaxing DSGE Restrictions

- Our prior for Φ^Δ has the property that its density is proportional to the expected likelihood ratio of $\Phi^* + \Phi^\Delta$ versus Φ^*.

- Likelihood ratio:

$$\begin{aligned} &\ln\left[\frac{\mathcal{L}(\Phi^* + \Phi^\Delta, \Sigma_u^* | Y_*, X_*)}{\mathcal{L}(\Phi^*, \Sigma_u^* | Y_*, X)}\right] \qquad (7) \\ &= -\frac{1}{2} tr\left[\Sigma_u^{*-1}\left(\Phi^{\Delta'} X_*' X_* \Phi^\Delta - 2\Phi^{*'} X_*' X_* \Phi^\Delta - 2(\Phi^* + \Phi^\Delta)' X_*' Y_* + 2\Phi^{*'} X_*' Y_*\right)\right]. \end{aligned}$$

- Taking expectations yields

$$I\!\!E_\theta^D\left[\ln\left[\frac{\mathcal{L}(\Phi^* + \Phi^\Delta, \Sigma_u^* | Y_*, X_*)}{\mathcal{L}(\Phi^*, \Sigma_u^* | Y_*, X_*)}\right]\right] = -\frac{1}{2} tr\left[\Sigma_u^{*-1}\left(\Phi^{\Delta'} \lambda T \Gamma_{XX} \Phi^\Delta\right)\right]. \qquad (8)$$

DSGE-VARs: Relaxing DSGE Restrictions

- We now choose a prior density that is proportional ($\propto$) to the expected likelihood ratio:

$$p(\Phi^{\Delta}|\Sigma_u^*) \propto \exp\left\{ -\frac{1}{2}tr\left[\lambda T \Sigma_u^{*-1}\left(\Phi^{\Delta'}\Gamma_{XX}\Phi^{\Delta}\right)\right]\right\}. \tag{9}$$

- Transform this prior into a prior for $\Phi = \Phi^*(\theta) + \Phi^{\Delta}$:

$$\Phi|\Sigma^*, \theta \sim \mathcal{N}\left(\Phi^*(\theta), \frac{1}{\lambda T}\left[\Sigma^{*-1} \otimes \Gamma_{XX}(\theta)\right]^{-1}\right). \tag{10}$$

- Relax the assumption that $\Sigma^{\Delta} = 0$.

- Again, we obtain:

$$\begin{aligned} \Sigma|\theta &\sim \mathcal{IW}\left(\lambda T \Sigma^*(\theta), \lambda T - k, n\right) \qquad (11)\\ \Phi|\Sigma, \theta &\sim \mathcal{N}\left(\Phi^*(\theta), \frac{1}{\lambda T}\left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta)\right]^{-1}\right), \end{aligned}$$

DSGE-VARs: Local Misspecification

- If we re-scale the misspecification as follows: $\Phi^{\Delta} = T^{-1/2}\widetilde{\Phi^{\Delta}}$, then the prior density becomes independent of the actual sample size:

$$p(\widetilde{\Phi^{\Delta}}|\Sigma^*, \theta) \propto \exp\left\{ -\frac{1}{2}tr\left[\lambda\Sigma^{*-1}\left(\widetilde{\Phi^{\Delta\prime}}\Gamma_{XX}(\theta)\widetilde{\Phi^{\Delta}}\right)\right]\right\} \tag{12}$$

- Large values of λ mean small misspecifications.

- $\widetilde{\Phi^{\Delta}}$ is "local" misspecification. DSGE model provides good albeit not perfect approximation to reality.

DSGE-VARs: Posteriors

- The joint posterior density of VAR and DSGE model parameters can be factorized:

$$p_\lambda(\Phi, \Sigma, \theta|Y) = p_\lambda(\Phi, \Sigma|Y, \theta) p_\lambda(\theta|Y). \tag{13}$$

The λ-subscript indicates the dependence of the posterior on the hyperparameter.

DSGE-VARs: Posteriors

- The posterior distribution of Φ and Σ is also of the Inverted Wishart – Normal form:

$$\begin{aligned}\Sigma|Y,\theta &\sim \mathcal{IW}\Big((1+\lambda)T\hat{\Sigma}_b(\theta),(1+\lambda)T-k,n\Big) \\ \Phi|Y,\Sigma,\theta &\sim \mathcal{N}\Big(\hat{\Phi}_b(\theta),\Sigma\otimes(\lambda T\Gamma_{XX}(\theta)+X'X)^{-1}\Big),\end{aligned} \tag{14}$$

- where $\hat{\Phi}_b(\theta)$ and $\hat{\Sigma}_b(\theta)$ are the given by

$$\begin{aligned}\hat{\Phi}_b(\theta) &= (\lambda T\Gamma_{XX}(\theta)+X'X)^{-1}(\lambda T\Gamma_{XY}+X'Y) \\ &= \left(\frac{\lambda}{1+\lambda}\Gamma_{XX}(\theta)+\frac{1}{1+\lambda}\frac{X'X}{T}\right)^{-1}\left(\frac{\lambda}{1+\lambda}\Gamma_{XY}+\frac{1}{1+\lambda}\frac{X'Y}{T}\right) \\ \hat{\Sigma}_b(\theta) &= \frac{1}{(1+\lambda)T}\Big[(\lambda T\Gamma_{YY}(\theta)+Y'Y)-(\lambda T\Gamma_{YX}(\theta)+Y'X) \\ &\quad\times(\lambda T\Gamma_{XX}(\theta)+X'X)^{-1}(\lambda T\Gamma_{XY}(\theta)+X'Y)\Big].\end{aligned}$$

DSGE-VARs: Posteriors

- The marginal posterior density of θ can be obtained by evaluating the marginal likelihood

$$
\begin{aligned}
p_\lambda(Y|\theta) \;=\; & \frac{|\lambda T \Gamma_{XX}(\theta) + X'X|^{-\frac{n}{2}} |(1+\lambda)T\hat{\Sigma}_b(\theta)|^{-\frac{(1+\lambda)T-k}{2}}}{|\lambda T \Gamma_{XX}(\theta)|^{-\frac{n}{2}} |\lambda T \Sigma^*(\theta)|^{-\frac{\lambda T-k}{2}}} \\
& \times \frac{(2\pi)^{-nT/2} 2^{\frac{n((1+\lambda)T-k)}{2}} \prod_{i=1}^n \Gamma[((1+\lambda)T-k+1-i)/2]}{2^{\frac{n(\lambda T-k)}{2}} \prod_{i=1}^n \Gamma[(\lambda T-k+1-i)/2]}.
\end{aligned}
$$

  and the prior density $p(\theta)$.

- We can also compute the marginal data density

$$
p_\lambda(Y) = \int p_\lambda(\theta|Y)p(\theta)d\theta. \tag{15}
$$

DSGE-VARs: Posteriors

MCMC Algorithm for DSGE-VAR:

1. Use the RWM Algorithm to generate draws $\theta^{(s)}$ from the marginal posterior distribution $p_\lambda(\theta|Y)$.

2. Use Geweke's modified harmonic mean estimator to obtain a numerical approximation of $\hat{p}_\lambda(Y)$.

3. For each draw $\theta^{(s)}$ generate a pair $\Phi^{(s)}$, $\Sigma^{(s)}$, by sampling from the $\mathcal{IW}-\mathcal{N}$ distribution.

DSGE-VARs: Posterior of θ

- Where does the information about θ come from? Rewrite posterior as

$$p(\Phi,\Sigma,\theta|Y) = p(\Phi,\Sigma|Y)p(\theta|\Phi,\Sigma). \tag{16}$$

  Projection of VAR estimates on DSGE model restriction.

- Consider quasi-likelihood function:

$$p^*(Y|\theta) \propto |\Sigma^*(\theta)|^{-T/2} \exp\left\{-\frac{1}{2} tr\left[{\Sigma^*}^{-1}(\theta)(Y - X\Phi^*(\theta))'(Y - X\Phi^*(\theta))\right]\right\}. \tag{17}$$

- Maximizing quasi-likelihood function with respect to θ is equivalent to minimizing the discrepancy between $\hat{\Phi}_{mle}$ and $\hat{\Sigma}_{mle}$ and the restriction functions $\Phi^*(\theta)$, $\Sigma^*(\theta)$.

$$\begin{aligned} \ln p^*(Y|\theta) &= -\frac{T}{2} vech(\hat{\Sigma}_{mle} - \Sigma^*(\theta))' D(\hat{\Sigma}_{mle}^{-1} \otimes \hat{\Sigma}_{mle}) D' vech(\hat{\Sigma}_{mle} - \Sigma^*(\theta))' \\ &\quad -\frac{1}{2} vec(\hat{\Phi}_{mle} - \Phi^*(\theta))'(\hat{\Sigma}_{mle}^{-1} \otimes X'X) vec(\hat{\Phi}_{mle} - \Phi^*(\theta)) \\ &\quad + const + small. \end{aligned} \tag{18}$$

DSGE-VARs: Posterior of θ

- *Proposition 1:* As $\lambda \longrightarrow \infty$ (weight of the prior tends to infinity), our procedure becomes equivalent to making inference based on the quasi-likelihood function $p^*(Y|\theta)$. Information accumulates at rate T.

- *Proposition 2:* As $\lambda \longrightarrow 0$, $T \longrightarrow \infty$, and $\lambda T \longrightarrow \infty$ (moderate weight of the prior, large sample), the marginal log-posterior density of θ is approximately quadratic in the discrepancy between $\hat{\Phi}_{mle}$ and $\hat{\Sigma}_{mle}$ and the restriction functions $\Phi^*(\theta)$, $\Sigma^*(\theta)$. Information accumulates at rate λT.

DSGE-VARs: Marginal Likelihood of λ

- We will study the fit of the DSGE model by examining the marginal likelihood function of the hyperparameter λ:

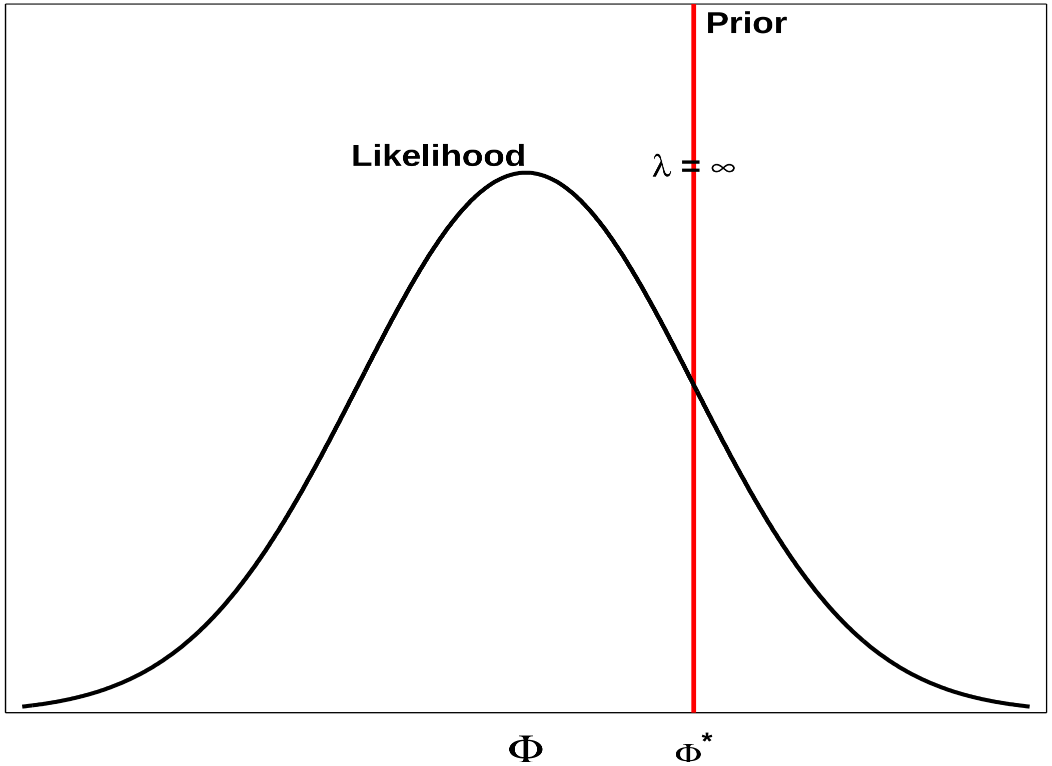
$$p(Y|\lambda) = \int p(Y|\theta, \Sigma, \Phi) p_\lambda(\theta, \Sigma, \Phi) d(\theta, \Sigma, \Phi). \tag{19}$$

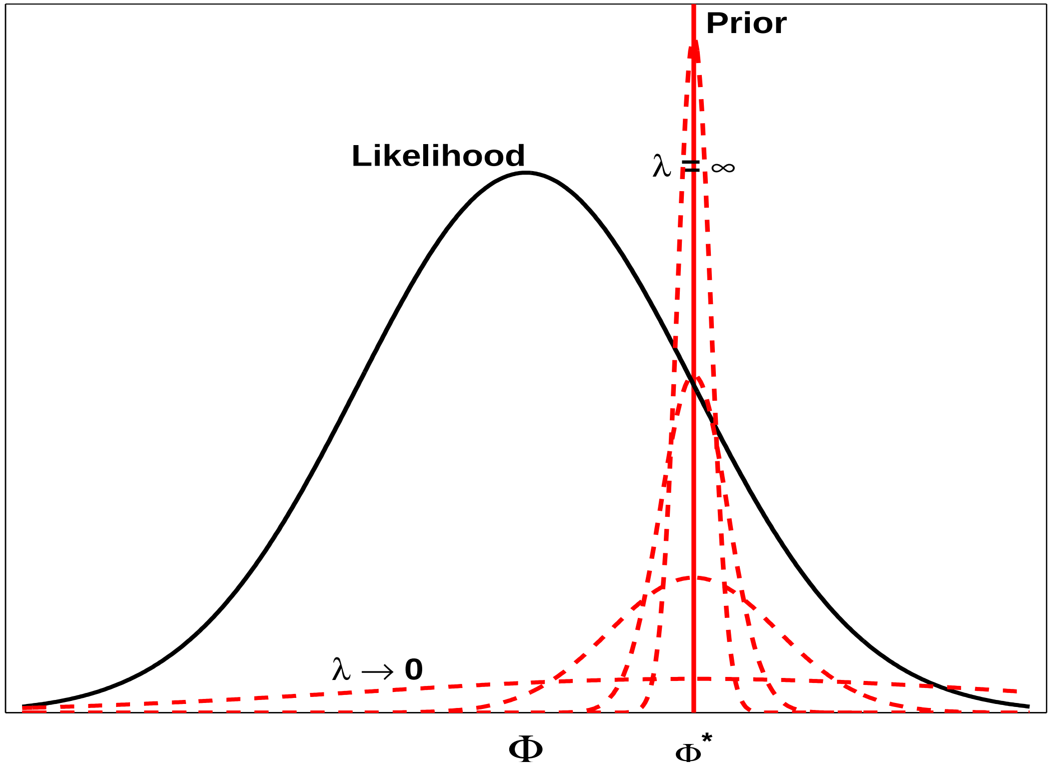
- Maximum / mode:

$$\hat{\lambda} = \mathrm{argmax}_{\lambda \in \Lambda}\, p(Y|\lambda).$$

- It is common in the literature to use marginal data densities to document the fit of DSGE models relative to VARs with diffuse priors. In our framework this corresponds to comparing

$$p(Y|\lambda = \text{small}) \quad \text{and} \quad p(Y|\lambda = \infty)$$


Prior
Likelihood
λ = ∞
Φ
Φ*


Prior
Likelihood
$\lambda = \infty$
$\lambda \to 0$
Φ
Φ^*

Marginal Likelihood of λ: Example

- Suppose the VAR takes the special form of an AR(1) model:

$$y_t = \phi y_{t-1} + u_t, \quad u_t \sim iid\mathcal{N}(0,1) \tag{20}$$

  and the DSGE model restricts ϕ to be equal to ϕ^*.

- Denote the DSGE model implied autocovariances by γ_0 and γ_1.

- Let $\hat{\gamma}_0$ and $\hat{\gamma}_1$ be sample autocovariances.

- Prior simplifies to

$$\phi \sim \mathcal{N}\left(\phi^*, \frac{1}{\lambda T \gamma_0}\right). \tag{21}$$

Marginal Likelihood of λ: Example

- Marginal likelihood of λ takes the following form

$$\ln p(Y|\lambda, \phi^*) = -T/2 \ln(2\pi) - \frac{T}{2}\tilde{\sigma}^2(\lambda, \phi^*) - \frac{1}{2}c(\lambda, \phi^*). \tag{22}$$

- The term $\tilde{\sigma}^2(\lambda, \phi^*)$ measures the in-sample one-step-ahead forecast error:

$$\lim_{\lambda \longrightarrow 0} \tilde{\sigma}^2(\lambda, \phi^*) = \frac{1}{T}\sum (y_t - \hat{\phi} y_{t-1})^2, \quad \lim_{\lambda \longrightarrow \infty} \tilde{\sigma}^2(\lambda, \phi^*) = \frac{1}{T}\sum (y_t - \phi^* y_{t-1})^2,$$

- The third term in (22) can be interpreted as a penalty for model complexity and is of the form

$$c(\lambda, \phi^*) = \ln\left(1 + \frac{\hat{\gamma}_0}{\lambda \gamma_0}\right).$$

- If an interior maximum of marginal likelihood exists, it is given by

$$\hat{\lambda} = \frac{\gamma_0 \hat{\gamma}_0^2}{T(\hat{\gamma}_0 \gamma_1 - \gamma_0 \hat{\gamma}_1)^2 - (\gamma_0)^2 \hat{\gamma}_0}. \tag{23}$$

Marginal Likelihood of λ: Example

- As λ approaches zero, the marginal log likelihood function tends to minus infinity.
- Consider the comparison of two models $\mathcal{M}_1$ ($\phi^*_{(1)}$) and $\mathcal{M}_2$ ($\phi^*_{(2)}$).
  - For small values of λ the goodness-of-fit terms are essentially identical. Marginal likelihoods differentials are due to differences in the penalty terms.
  - For large values of λ, marginal likelihood comparison is driven by the relative in-sample fit of the two restricted specifications.

DSGE-VARs: Posterior of λ

- Numerical Illustration in An and Schorfheide (2005):

Specification	Data Set 1	Data Set 2
DSGE Model	-196.66	-279.38
DSGE-VAR $\lambda = \infty$	-196.88	-277.49
DSGE-VAR $\lambda = 5.00$	-198.87	-270.46
DSGE-VAR $\lambda = 1.00$	-206.57	-258.25
DSGE-VAR $\lambda = 0.75$	-209.53	-257.53
DSGE-VAR $\lambda = 0.50$	-215.06	-258.73
DSGE-VAR $\lambda = 0.25$	-231.20	-269.66

DSGE-VARs: Comparison of DSGE and VAR

- Goal of IRF comparisons is to document in which dimensions the DSGE model dynamics are (in)consistent with the data.
- To what extent does the VAR satisfy key structural equations implied by the DSGE? E.g., is the Phillips curve equation misspecified?
- Examples: Cogley and Nason (1994), Rotemberg and Woodford (1997), Schorfheide (2000), Boivin and Giannoni (2003), and Christiano, Eichenbaum, and Evans (2004), to name a few.
- Important issue: estimation and the identification of the VAR that serves as a benchmark. Problems: many parameters to estimate, many shocks to identify.

DSGE-VARs: Comparison of DSGE and VAR

- In our framework: compare (i) DSGE-VAR(∞) and DSGE-VAR($\hat{\lambda}$) IRFs; (ii) DSGE-VAR($\hat{\lambda}$) and DSGE IRFs.

DSGE-VARs: Identification

- So far the DSGE-VAR is reduced form. For most applications we would like a mapping from VAR innovations into structural shocks.

- An (exactly) identified VAR is a triplet: (Φ, Σ, Ω), where Ω is orthonormal:

$$\left(\frac{\partial y_t}{\partial \epsilon_t'} \right)_{VAR} = \Sigma_{tr} \Omega.$$

- The DSGE model is identified: there is a matrix $\Omega^*(\theta)$ that maps the variance-covariance matrix of innovations into the portion attributed to each shock:

$$\left(\frac{\partial y_t}{\partial \epsilon_t'} \right)_{DSGE} = \Sigma_{tr}^*(\theta) \Omega^*(\theta).$$

- Identified DSGE-VAR: $(\Phi, \Sigma, \Omega^*(\theta))$.

DSGE-VARs: Identification

MCMC Algorithm for DSGE-VAR:

1. Use the RWM Algorithm to generate draws $\theta^{(s)}$ from the marginal posterior distribution $p_\lambda(\theta|Y)$.

2. Use Geweke's modified harmonic mean estimator to obtain a numerical approximation of $\hat{p}_\lambda(Y)$.

3. For each draw $\theta^{(s)}$ generate a pair $\Phi^{(s)}$, $\Sigma^{(s)}$, by sampling from the $\mathcal{IW}-\mathcal{N}$ distribution. Moreover, compute the orthonormal matrix $\Omega^{(s)}$ as described above.

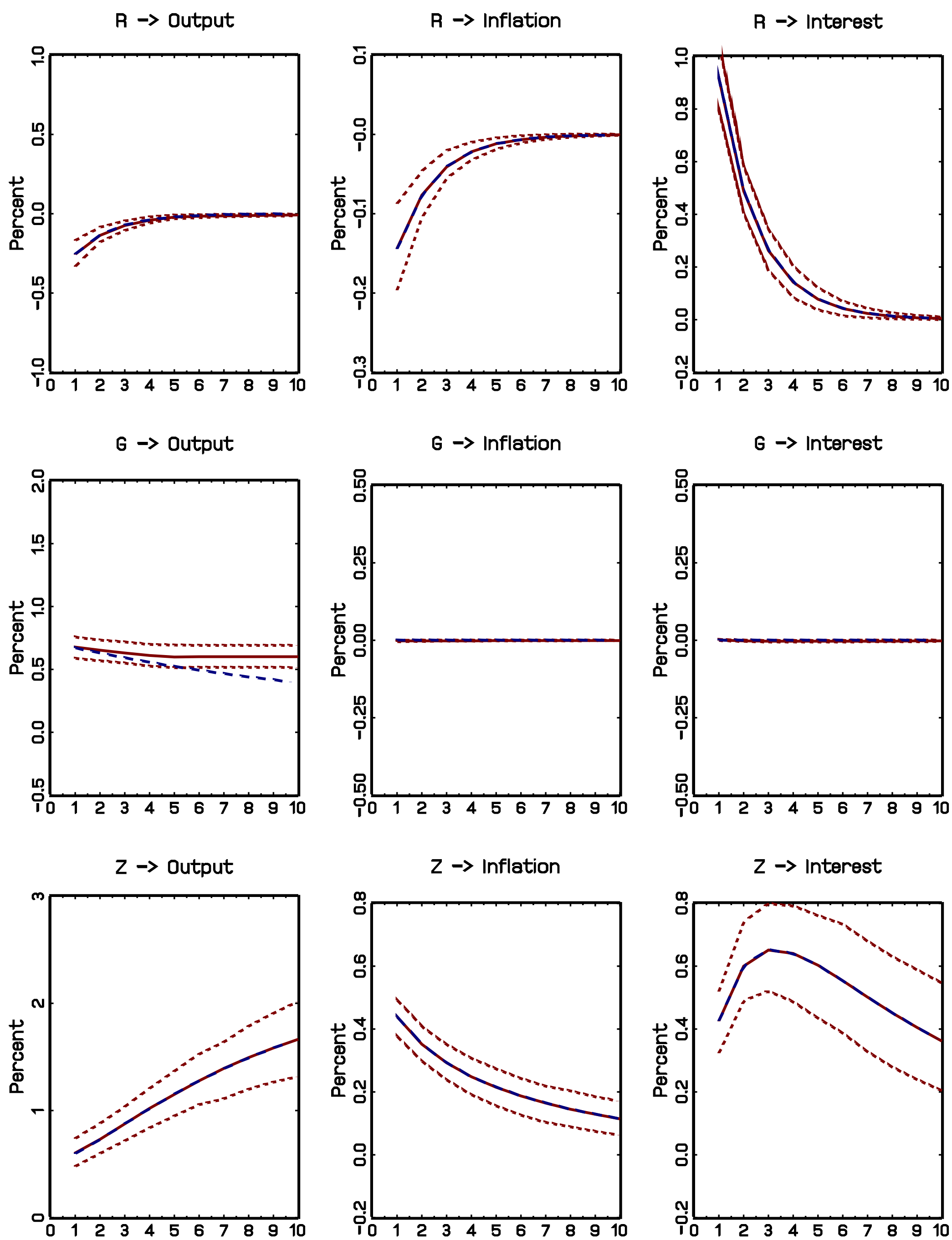
DSGE-VARs: IRF Comparisons

- How well is the state-space representation of the linearized DSGE model approximated by the finite-order VAR?

- For each θ draw compare responses of the state-space version of the DSGE to the DSGE-VAR($\lambda = \infty$) version.

(insert figures here)

Figure 11: Impulse Responses, DSGE and DSGE-VAR($\lambda = \infty$) – Model $\mathcal{M}_1(L)$, Data $\mathcal{D}_5(L)$



Notes: DSGE model responses computed from state-space representation: posterior mean (solid); DSGE-VAR($\lambda = \infty$) responses: posterior mean (dashed) and pointwise 90% probability bands (dotted).

DSGE-VARs: IRF Comparisons

- How different are the IRFs of the VAR that is estimated subject to the DSGE model restrictions from the IRFs of the VAR in which restrictions are relaxed?

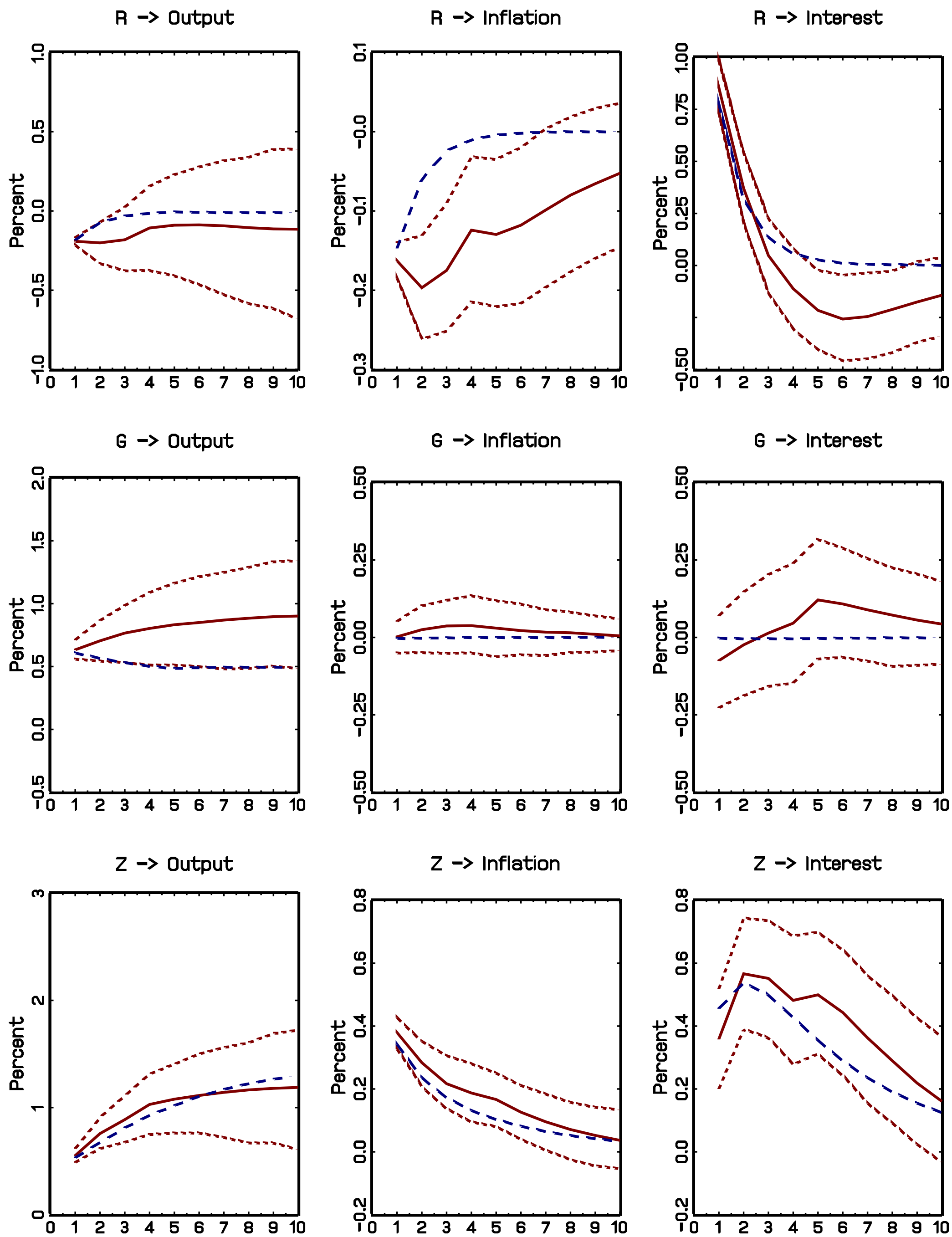
- For each (Φ, Σ, θ) draw compare responses of the state-space version of the DSGE to the DSGE-VAR($\lambda = \infty$) version.

- We plot posterior mean responses of DSGE-VAR($\lambda = \infty$).

- Moreover, for each draw we compute the difference between DSGE-VAR(λ) and DSGE-VAR($\lambda = \infty$). We use these differences to compute a posterior mean and 90% probability bands.

(insert figures here)

Figure 12: Impulse Responses, DSGE-VAR($\lambda = \infty$) and DSGE-VAR($\lambda = 1$) – Model $\mathcal{M}_1(L)$, Data $\mathcal{D}_5(L)$



Notes: DSGE-VAR($\lambda = \infty$) posterior mean responses (solid), DSGE-VAR($\lambda = 1$) posterior mean responses (long dashes). Pointwise 90% probability bands (short dashes) signify shifted probability intervals for the difference between $\lambda = \infty$ and $\lambda = 1$ responses.

DSGE-VARs: IRF Comparisons

- Suppose we rewrite the structural equations as follows:

$$
\begin{aligned}
\hat{y}_t - \hat{y}_{t+1} + \frac{1}{\tau}[\hat{R}_t - \hat{\mathbb{E}}_t \pi_{t+1}] &= (1-\rho_g)\hat{g}_t + \mathbb{E}_t \hat{z}_{t+1} \\
\hat{\pi}_t - \beta \mathbb{E}_t[\hat{\pi}_{t+1}] - \kappa \hat{y}_t &= -\kappa \hat{g}_t \\
\hat{R}_t - \rho_R \hat{R}_{t-1} - (1-\rho_R)\psi_1 \hat{\pi}_t + (1-\rho_R)\psi_2 \hat{y}_t &= -(1-\rho_R)\psi_2 \hat{g}_t + \epsilon_{R,t}
\end{aligned}
$$

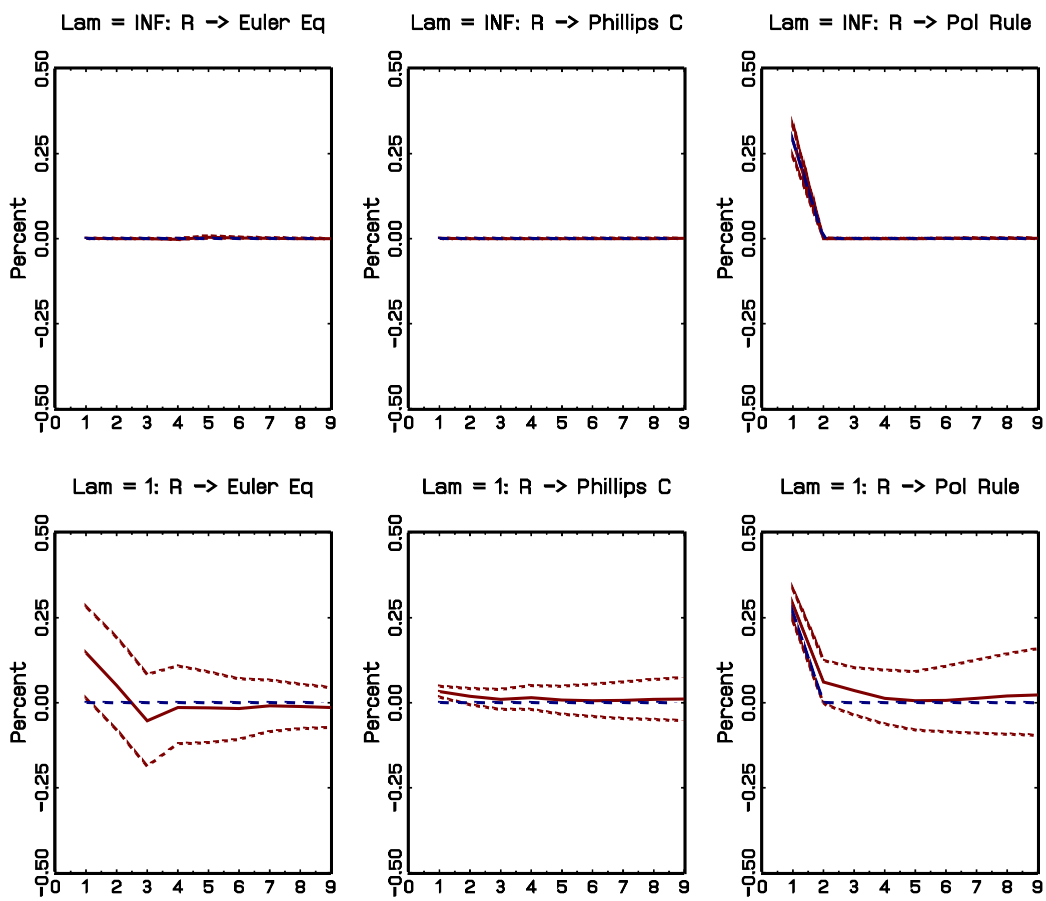
- For instance, in response to a monetary policy shock, the right-hand-side of the Euler equation and the Phillips curve equation has to be zero.

- We can check these conditions for the DSGE-VAR(λ) response.

- We overlay the right-hand-side for DSGE and DSGE-VAR.

(insert figures here)

Figure 13: Impulse Responses, DSGE-VAR($\lambda = \infty$) and DSGE-VAR($\lambda = 1$) – Model $\mathcal{M}_1(L)$, Data $\mathcal{D}_5(L)$



Notes: DSGE model responses: posterior mean (solid); DSGE-VAR responses: posterior mean (dashed) and pointwise 90% probability bands (dotted).

DSGE-VARs: Some Empirical Results

- Based on Del Negro, Schorfheide, Smets, and Wouters (2006)...
- U.S. data; unless noted otherwise thirty years of observations ($T = 120$), starting in QII:1974 and ending in QI:2004. Lag length $p = 4$.
- Forecasting exercise: beginning from QIII:1954 we construct 58 rolling samples of 120 observations.
- Four parts of the analysis:
  - Parameters: priors and posteriors
  - Marginal likelihood function
  - Model comparison: baseline versus *No Habit* and *No Indexation.*
  - Some pseudo-out-of-sample forecast error statistics.

DSGE-VARs: Some Empirical Results

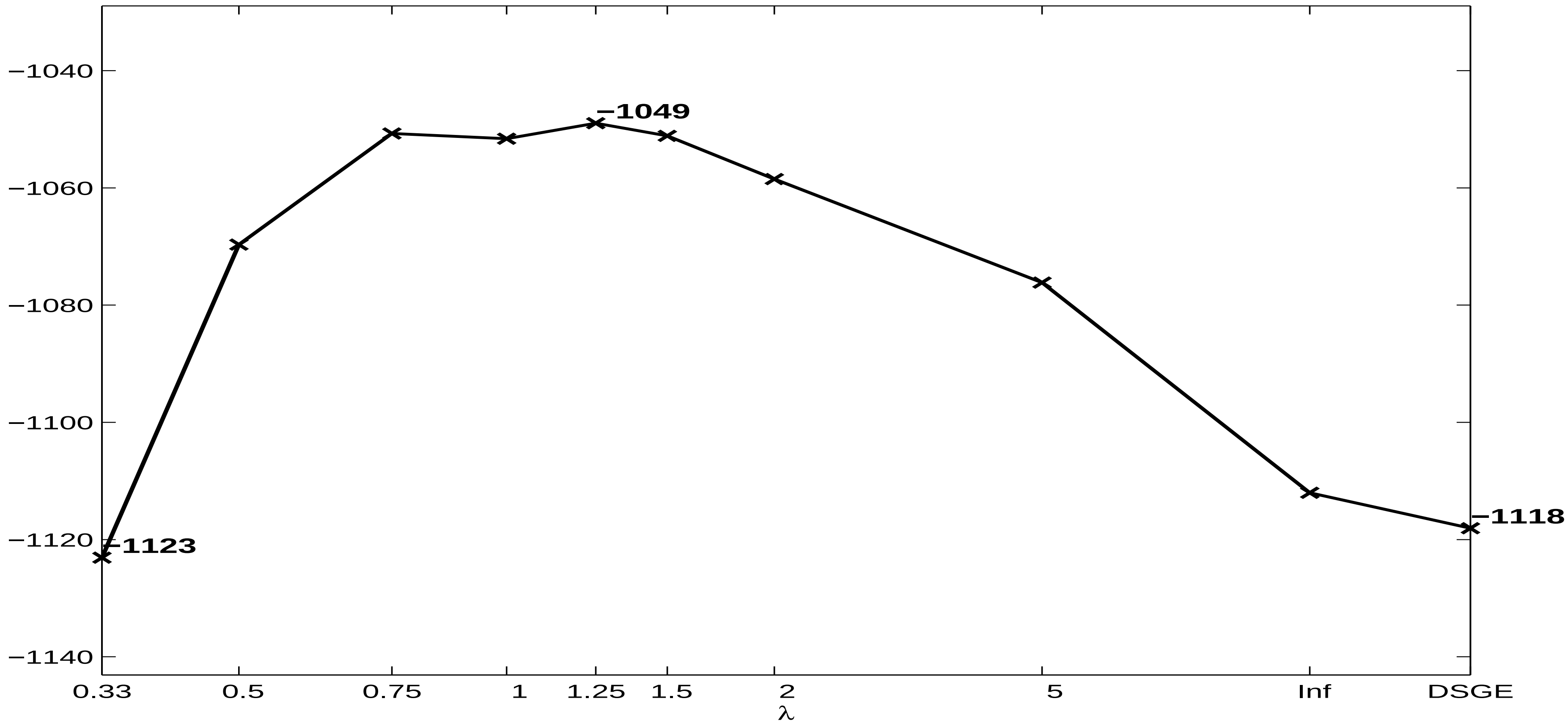
	Prior		DSGE-VECM($\hat{\lambda}$) Post.		DSGE Post.	
	Mean	Interval	Mean	Interval	Mean	Interval
ζ_p	0.60	[0.29 , 0.93]	0.79	[0.72 , 0.86]	0.83	[0.79 , 0.87]
ι_p	0.50	[0.08 , 0.95]	0.75	[0.53 , 1.00]	0.76	[0.57 , 0.97]
ζ_w	0.60	[0.29 , 0.94]	0.79	[0.70 , 0.87]	0.89	[0.84 , 0.93]
ι_w	0.50	[0.05 , 0.93]	0.45	[0.04 , 0.80]	0.70	[0.47 , 0.96]
h	0.70	[0.62 , 0.78]	0.75	[0.70 , 0.81]	0.81	[0.77 , 0.85]
ψ_1	1.50	[0.99 , 2.09]	1.80	[1.42 , 2.19]	2.21	[1.79 , 2.63]
ψ_2	0.20	[0.05 , 0.35]	0.16	[0.09 , 0.22]	0.07	[0.03 , 0.10]
ρ_r	0.50	[0.18 , 0.83]	0.76	[0.70 , 0.83]	0.82	[0.78 , 0.86]

DSGE-VARs: Some Empirical Results

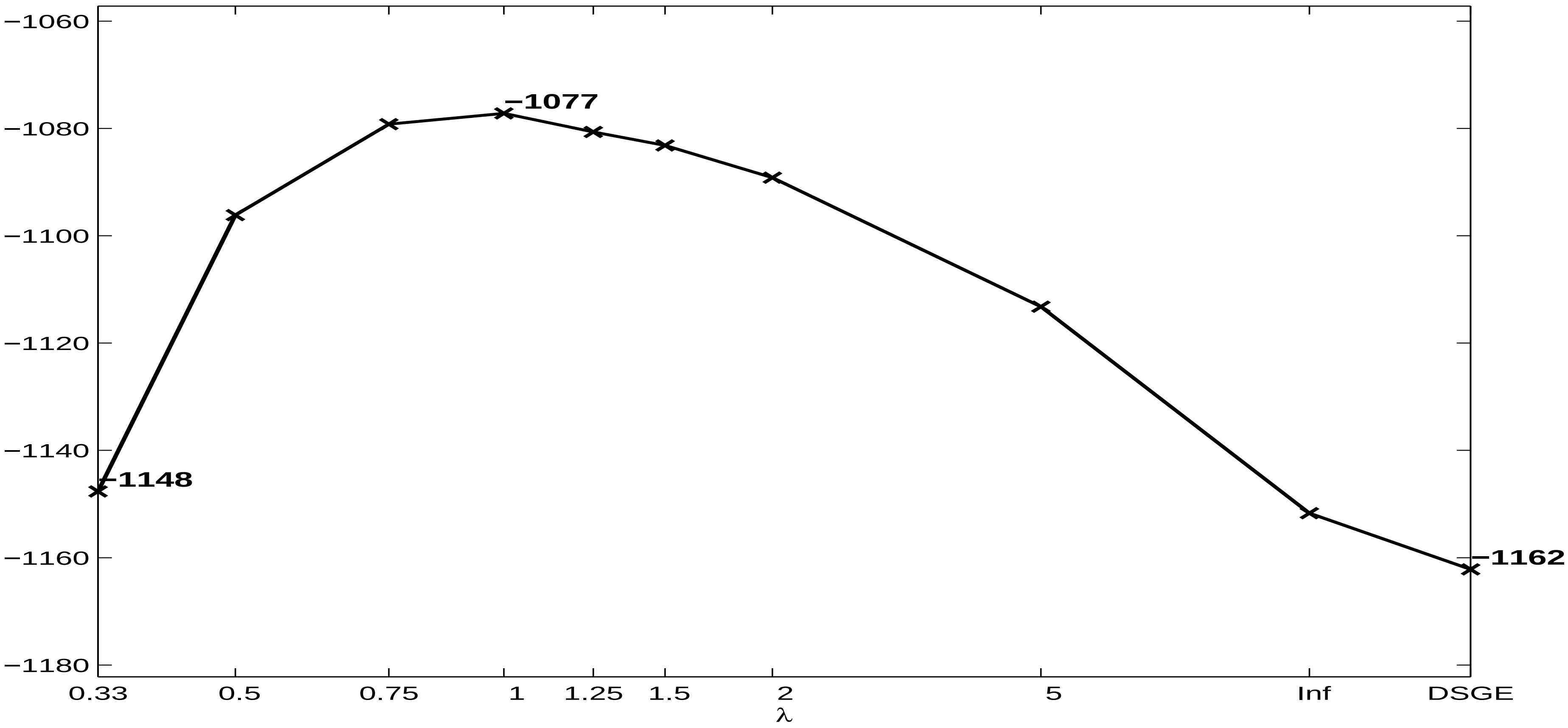
- Roughly: $\lambda = \infty$ dogmatically imposes DSGE model restrictions; $\lambda = 0$ completely ignores the restrictions.

- Best fit in terms of Bayesian marginal likelihood and out-of-sample forecasting performance is obtained for intermediate values of λ – indicating some disagreement between DSGE and data.

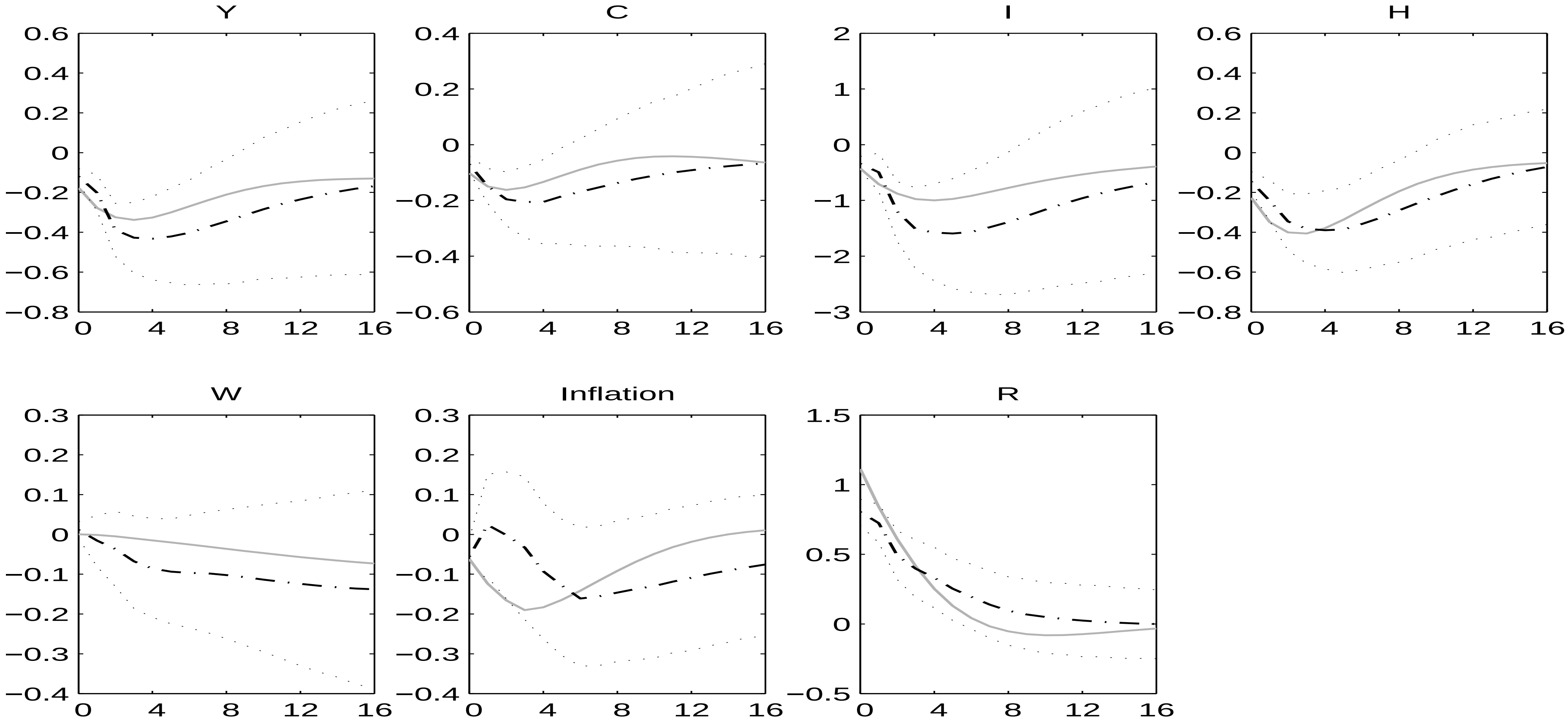
Marginal Likelihood of λ: 30-Year Sample: QII:1974 to QI:2004



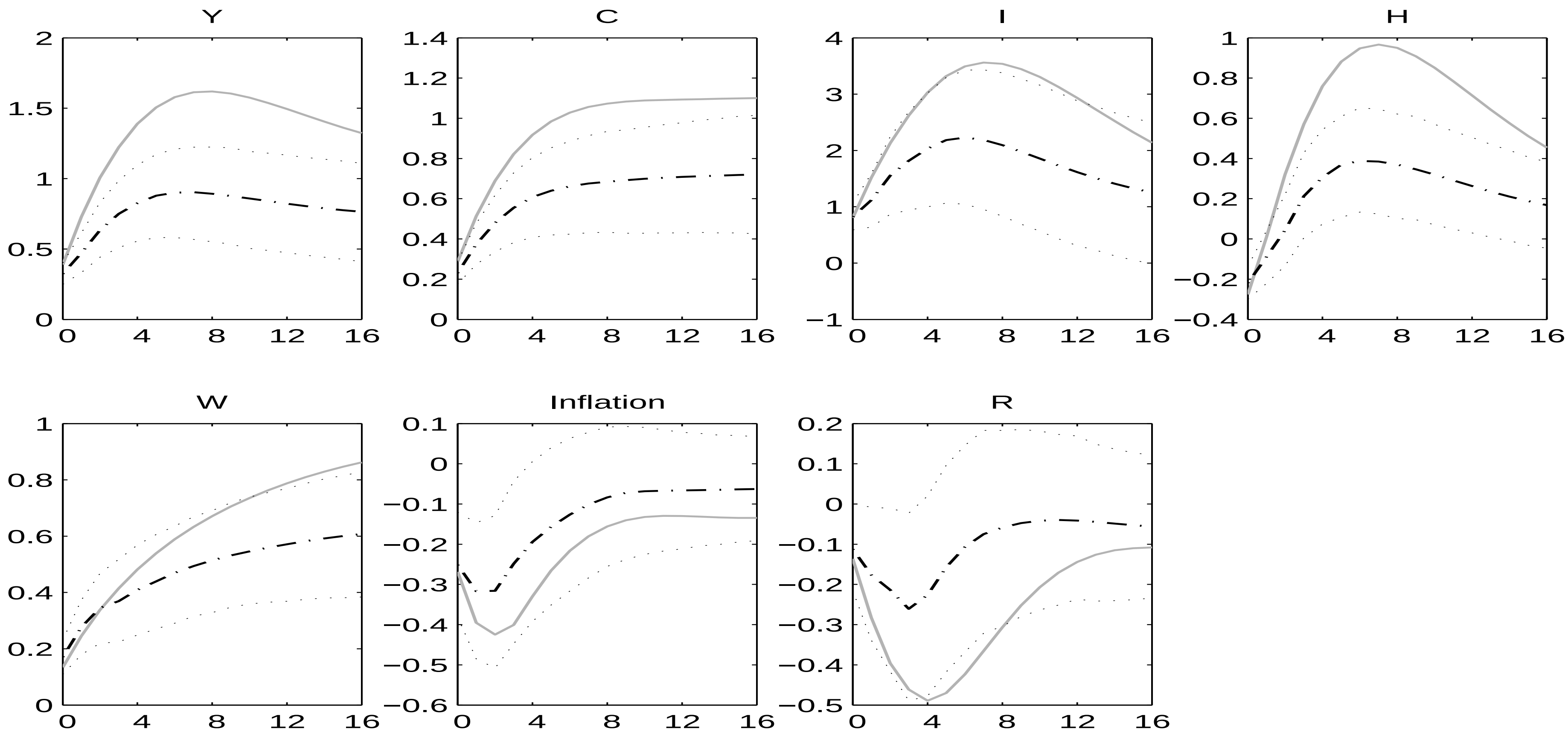
Marginal Likelihood of λ: 30-Year Sample: QII:1970 to QI:2000



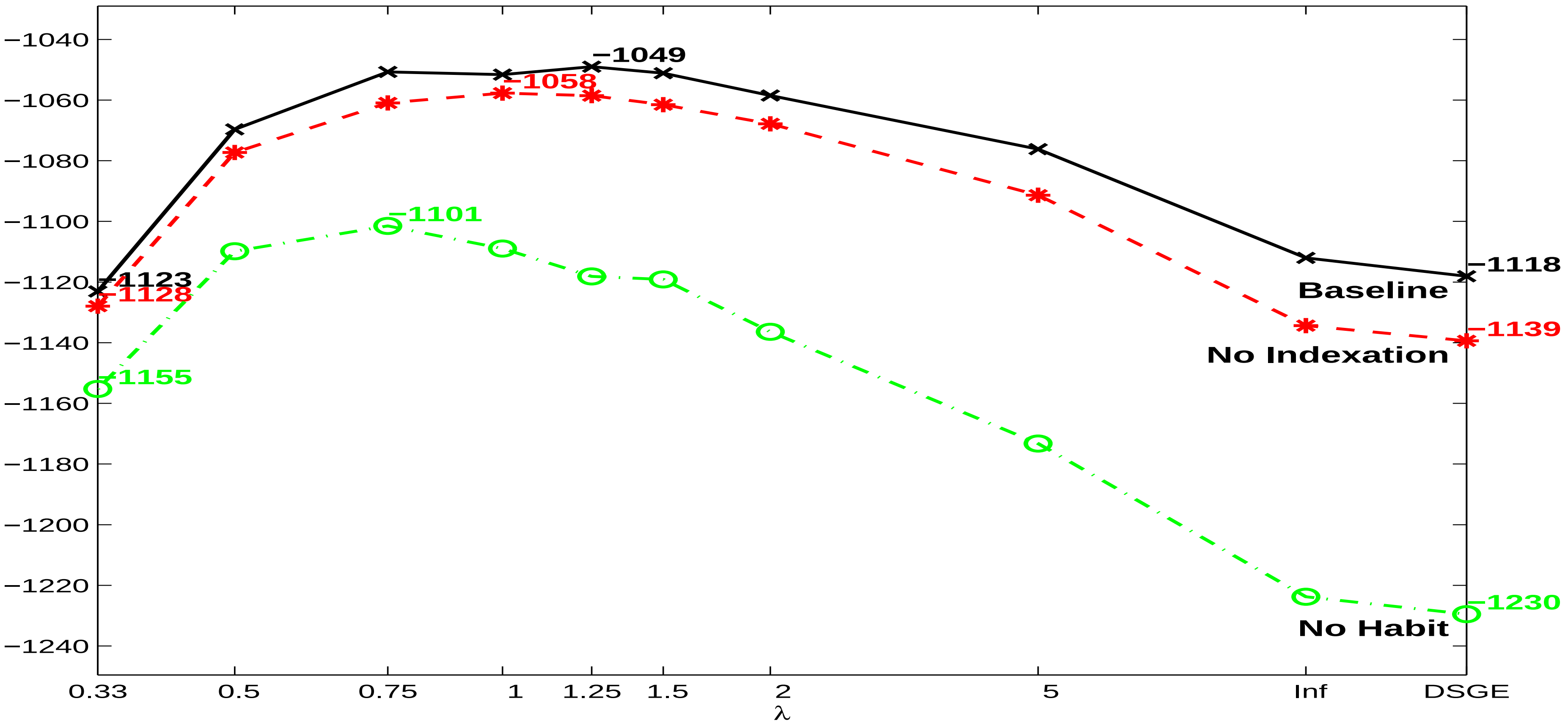
Monetary Policy Shock IRFs: DSGE-VAR($\hat{\lambda}$) vs. DSGE-VAR(∞)



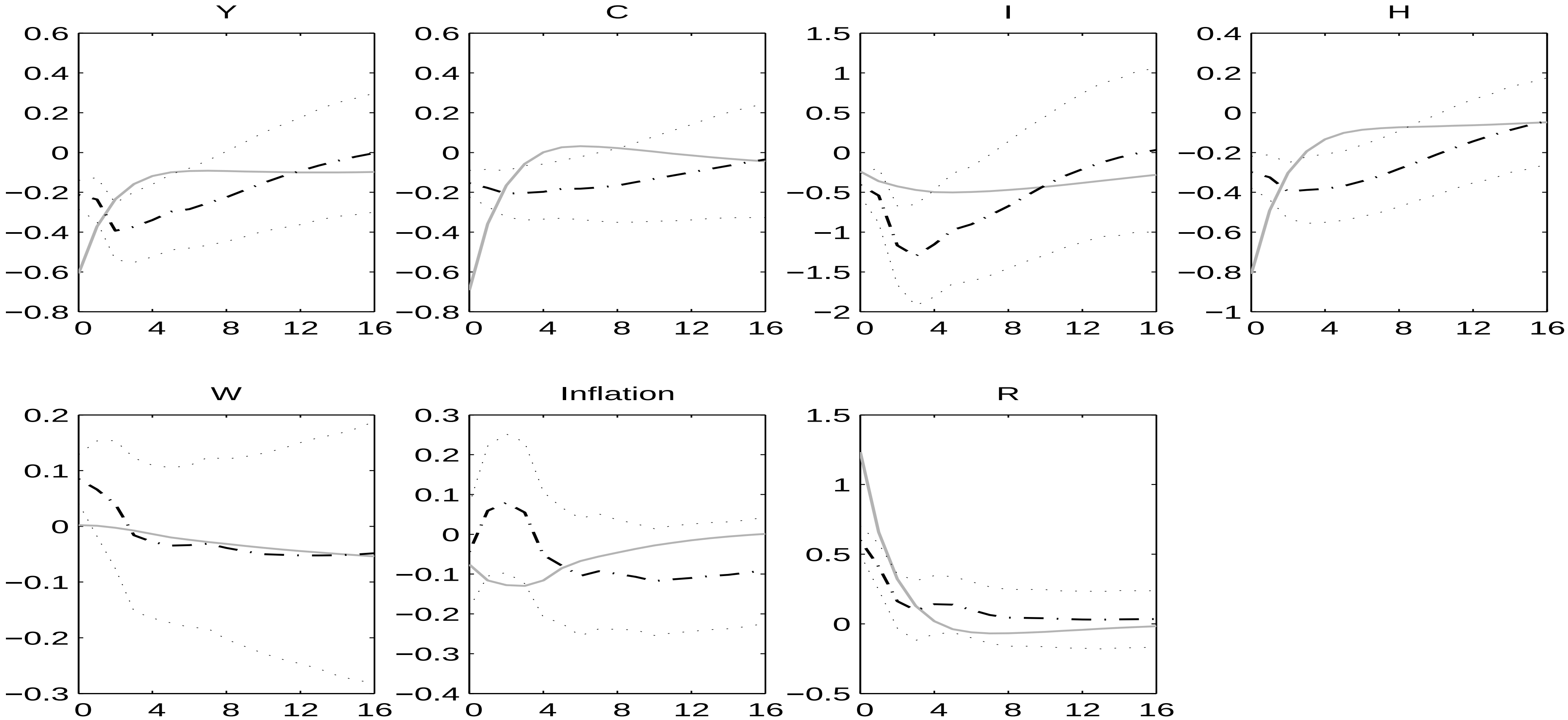
Technology Shock IRFs: DSGE-VAR($\hat{\lambda}$) vs. DSGE-VAR(∞)



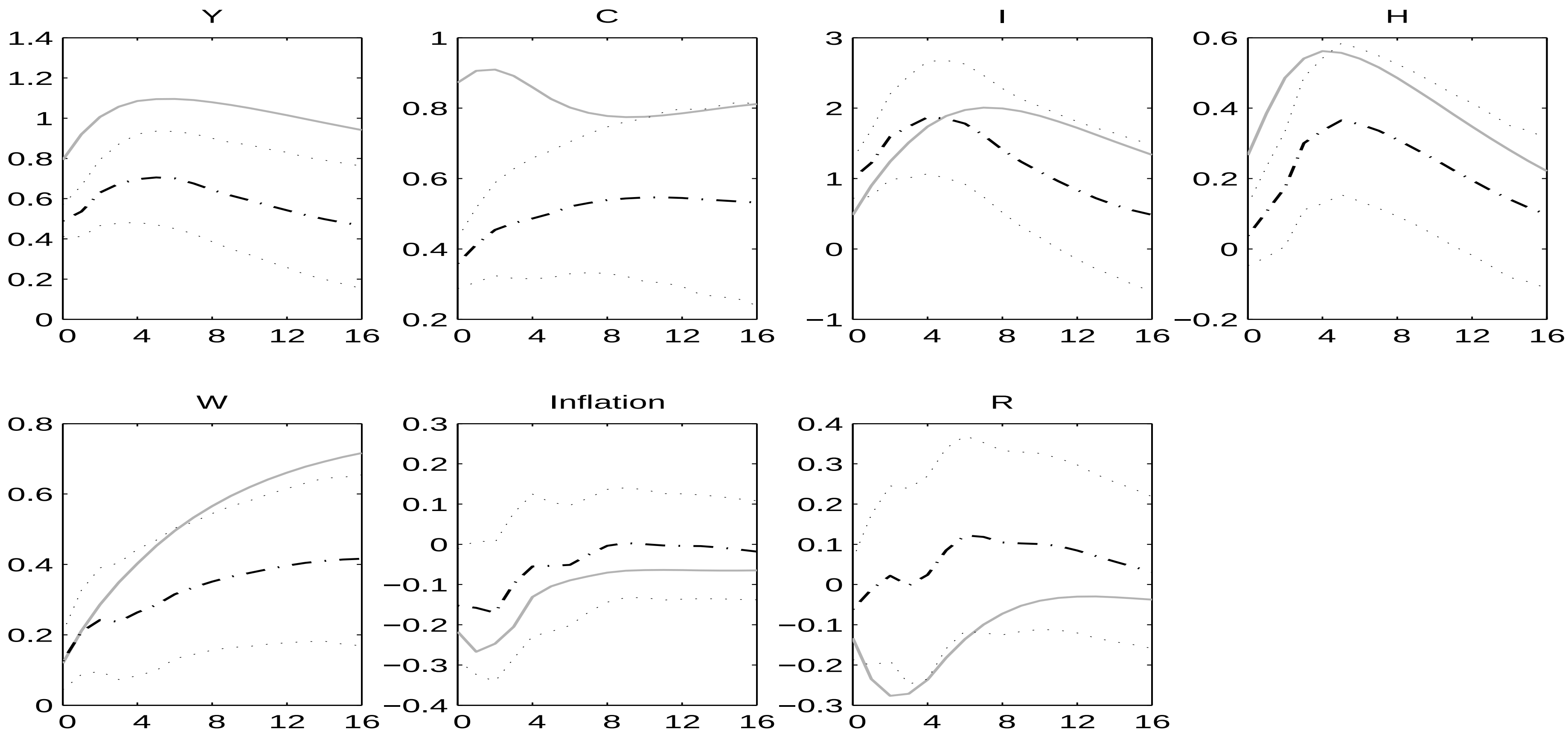
Comparing Different DSGE Model Specifications



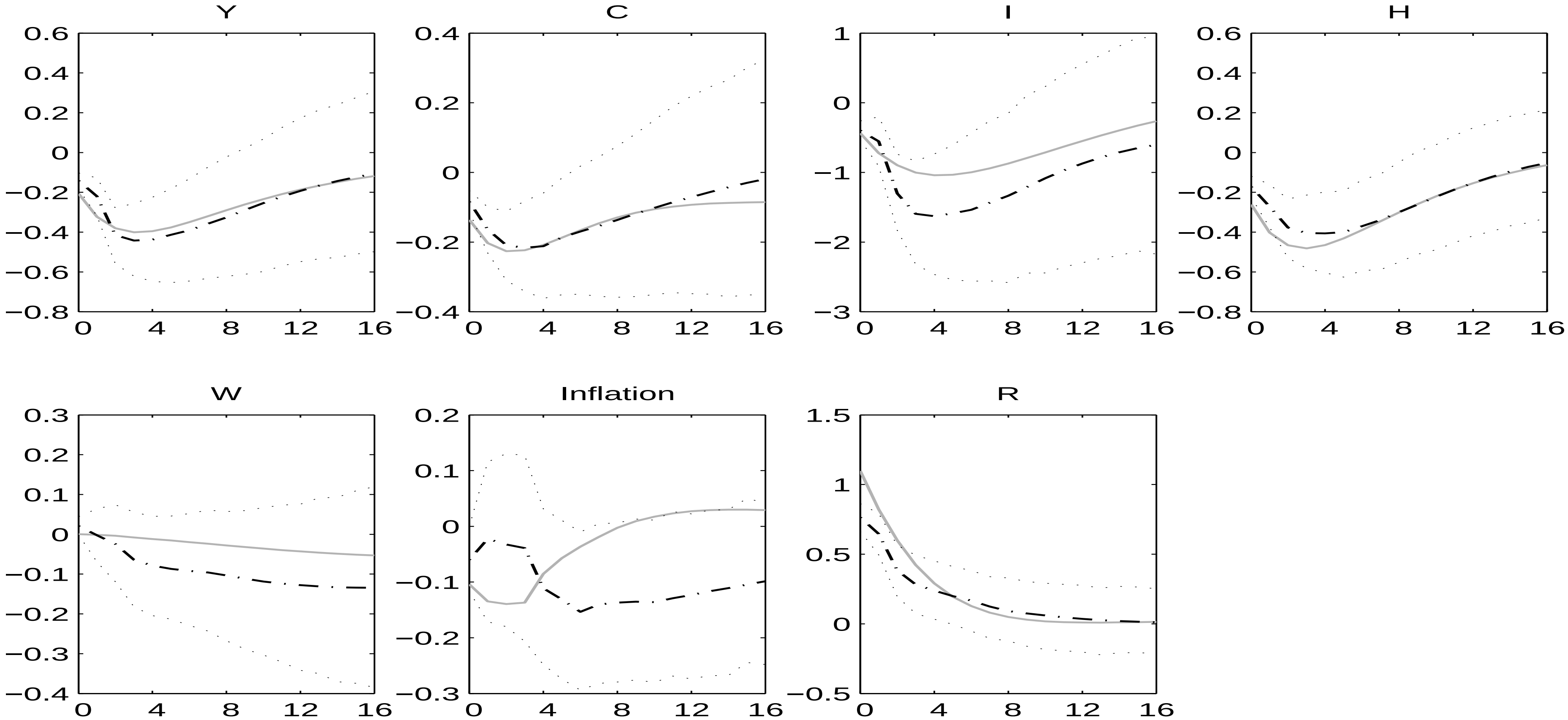
No Habit Specification: Monetary Policy Shock IRFs



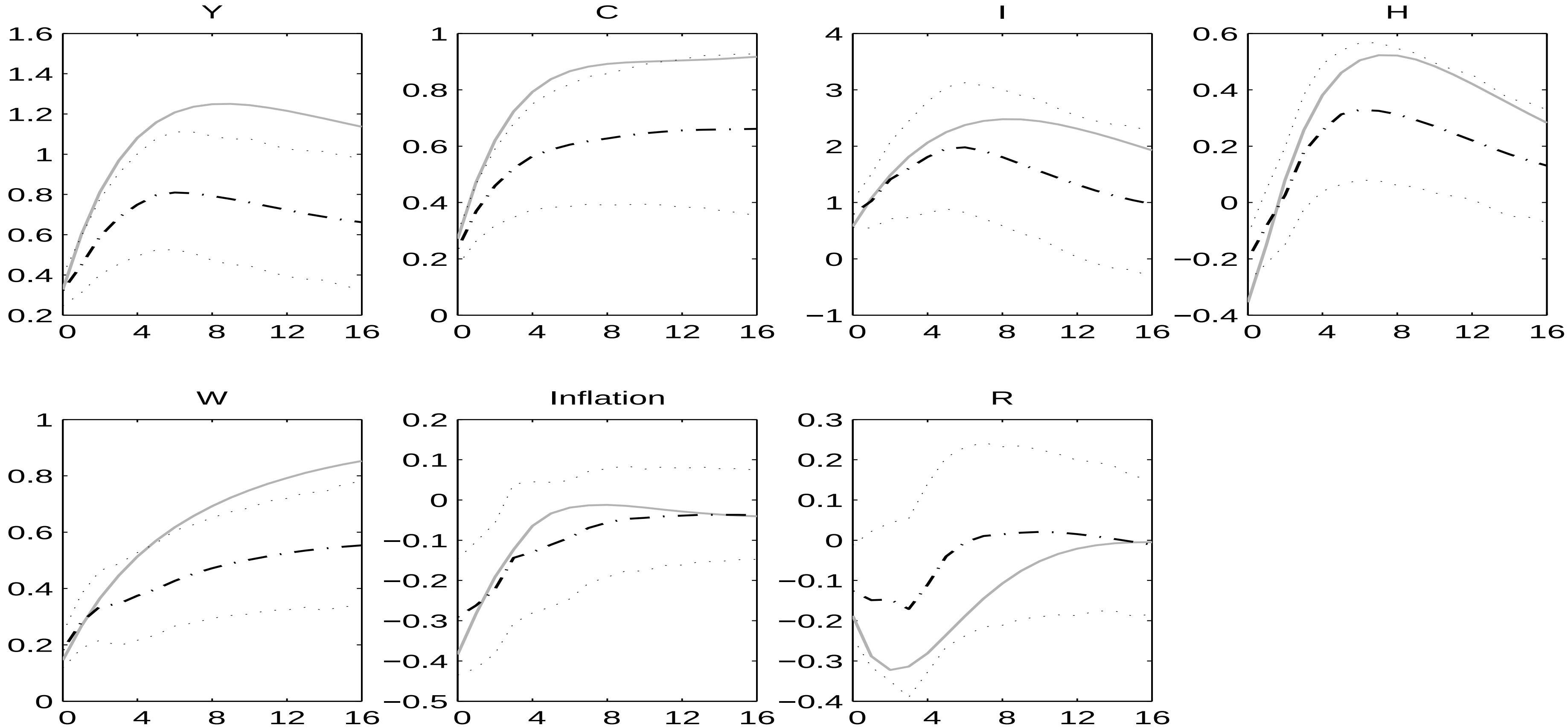
No Habit Specification: Technology Shock IRFs



No Indexation Specification: Monetary Policy Shock IRFs



No Indexation Specification: Technology Shock IRFs

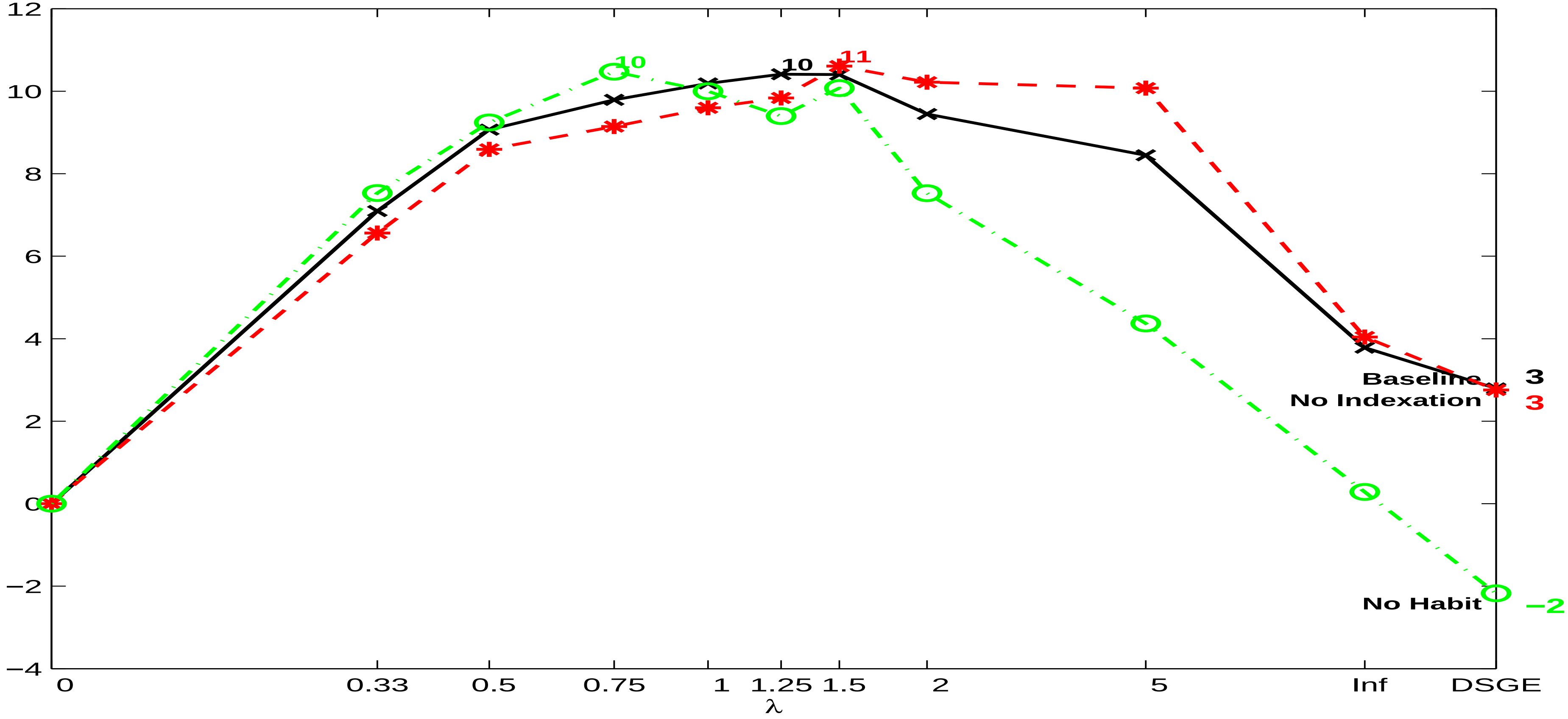


Discussion

Question: If we did not have the baseline model at hand, but only the alternative specification, can we learn something from our procedure about what is missing?

- *No Habit*: For money shock (but also technology), clearly something is amiss! Consumption (Y, and H) responds to quickly in DSGE-VAR(∞) relative to DSGE-VAR($\hat{\lambda}$
- ...however DSGE-VAR($\hat{\lambda}$)'s IRFs are close to those of the Baseline model: Even if the DSGE model w/o Habit is grossly misspecified, the DSGE-VAR($\hat{\lambda}$) is not too bad as a benchmark.
- *No Indexation*: No clear evidence from Money and/or Tech IRFs that a feature that can lead to improved fit is missing.

Forecasting Performance: One-period Ahead Root Mean Square Error Summary



Discussion

- Roughly: DSGE model and unrestricted VAR are comparable. DSGE-VAR improves.
- This suggests:
  - Unrestricted VARs are not a good benchmark for DSGE evaluations.
  - DSGE-VARs are useful tools for central banks

DSGE-VARs: Extensions

- Start from DSGE with interest-rate feedback rule, allow for deviations from cross-coefficient restrictions while maintain form of policy rule – leads to a collection of identified VARs.

- Observables y_t: Interest Rate, Inflation, Output Gap.

- Let $x_t' = [y_{t-1}', \ldots, y_{t-p}', 1]$. Rewrite policy rule in general terms:

$$\underbrace{y_{1,t}}_{Interest\ Rate} = x_t' M_1 \beta_1(\theta) + y_{2,t}' M_2 \beta_2(\theta) + \epsilon_{1,t}. \tag{24}$$

- Define $\Psi^*(\theta) = \left(I\!\!E_\theta^D [x_t x_t'] \right)^{-1} I\!\!E_\theta^D [x_t y_{2,t}']$ and write remainder of system:

$$\underbrace{y_{2,t}'}_{Inflation,\ OutputGap} = x_t' \Psi^*(\theta) + u_{2,t}'. \tag{25}$$

- VAR approximation (25) is in general not exact yet quite accurate with four lags in our application.

- (24) and (25) comprise a partially identified (based on exclusion restrictions) VAR.

DSGE-VARs: Extensions

- Rewrite Interest Rate equation

$$y_{1,t} = x_t' M_1 \beta_1(\theta) + x_t' \Psi^*(\theta) M_2 \beta_2(\theta) + u_{1,t}, \tag{26}$$

- and create restricted VAR for y_t

$$y_t' = x_t' \Phi + u_t', \quad I\!\!E[u_t u_t'] = \Sigma \tag{27}$$

with

$$\Phi = \Phi^*(\theta) = B_1(\theta) + \Psi^*(\theta) B_2(\theta), \quad \Sigma = \Sigma^*(\theta).$$

DSGE-VARs: Extensions

- There is a vector θ and matrices Ψ^Δ and Σ^Δ such that the data are generated from the VAR in Eq. (27)

$$\Phi = B_1(\theta) + (\Psi^*(\theta) + \Psi^\Delta)B_2(\theta), \quad \Sigma = \Sigma^*(\theta) + \Sigma^\Delta. \tag{28}$$

- (Assume $\Sigma^\Delta = 0$) Construct a prior with property that its density is proportional to the expected likelihood ratio of Ψ evaluated at its (misspecified) restricted value $\Psi^*(\theta)$ versus the "true" value $\Psi = \Psi^*(\theta) + \Psi^\Delta$.

- Same analysis as before... MCMC is a bit more complicated.

DSGE-VARs: Extensions

- The forecast error $u_{2,t}$ is a function of the structural shocks: $u'_{2,t} = \epsilon_{1,t} A_1 + \epsilon'_{2,t} A_2$.
- After some matrix algebra we can determine A_1 and $A_2' A_2$, which identifies monetary policy shocks, but does not separate technology from demand shocks.
- We follow the idea in Del Negro and Schorfheide (2004) and decompose the DSGE model response $A_2^{D'}(\theta) = A_{2,tr}^{D'}(\theta)\Omega^*(\theta)$.
- And then let: $A_2 = chol(A_2' A_2)\Omega^*(\theta)$.
- Now we have a collection of identified VARs. Del Negro and Schorfheide (2005) study effects of changes in the policy rule in this framework.

Conclusions and Outlook

- Large body of empirical work on the Bayesian estimation / evaluation of DSGE models.
- An and Schorfheide (2005) paper illustrates techniques based on New Keynesian model.
- Model size / dimensionality of the parameter space pose challenge for MCMC methods.
- Lack of identification of structural parameters is often difficult to detect; creates a challenge for scientific reporting.
- Model misspecification is and will remain a concern in empirical work with DSGE models despite continuous efforts by macroeconomists to develop more adequate models.
- Hence it is important to develop methods that incorporate potential model misspecification in the measures of uncertainty constructed for forecasts and policy recommendations.