

Comments on “On the Fit of Forecasting
Performance of New-Keynesian Models” by Del
Negro, Schorfheide, Smets, and Wouters

James D. Hamilton
UCSD

Data:

- 1) output
- 2) consumption
- 3) investment
- 4) employment
- 5) real wage
- 6) inflation
- 7) nominal interest rate

Shocks:

- 1) productivity
- 2) intertemporal substitution
- 3) relative cost of investment goods
- 4) consumption-leisure substitution
- 5) fiscal shock (fraction to govt)
- 6) mark-up shock
- 7) Fed policy

θ vector of truly structural parameters

(discount rate, serial dependence of productivity shock, capital share,...)

y_t vector of data (growth rates, deviations from mean,...)

z_t cointegrating variables

$(\ln C_t - \ln Y_t, \ln I_t - \ln Y_t, \ln(W_t/P_t) - \ln Y_t)'$

DSGE:

$$\mathbf{y}_t = \mathbf{c}(\boldsymbol{\theta}) + \mathbf{B}(\boldsymbol{\theta})\mathbf{z}_{t-1} + \boldsymbol{\Phi}_1(\boldsymbol{\theta})\mathbf{y}_{t-1} \\ + \boldsymbol{\Phi}_2(\boldsymbol{\theta})\mathbf{y}_{t-2} + \cdots + \boldsymbol{\Phi}_p(\boldsymbol{\theta})\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

$$E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega}(\boldsymbol{\theta})$$

VAR:

$$\mathbf{y}_t = \mathbf{c} + \mathbf{B}\mathbf{z}_{t-1} + \boldsymbol{\Phi}_1\mathbf{y}_{t-1} + \boldsymbol{\Phi}_2\mathbf{y}_{t-2} \\ + \cdots + \boldsymbol{\Phi}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

$$E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega}$$

$\boldsymbol{\theta}$ has much fewer elements than VAR

$$\mathbf{A} = \begin{bmatrix} \mathbf{c} & \mathbf{B} & \Phi_1 & \Phi_2 & \cdots & \Phi_p \end{bmatrix}$$

$$\mathbf{x}_t = (1, \mathbf{z}'_t, \mathbf{y}'_{t-1}, \mathbf{y}'_{t-2}, \dots, \mathbf{y}'_{t-p})'$$

$$\mathbf{y}_t = \mathbf{A}\mathbf{x}_t + \boldsymbol{\varepsilon}_t$$

$$\mathbf{a} = \text{vec}(\mathbf{A})$$

$$\mathbf{a} | \boldsymbol{\theta}, \boldsymbol{\Omega}, \lambda \sim N(\mathbf{a}(\boldsymbol{\theta}), (\lambda T)^{-1} [\boldsymbol{\Omega} \otimes \boldsymbol{\Gamma}_{\mathbf{xx}}(\boldsymbol{\theta})^{-1}])$$

$$\mathbf{a}|\boldsymbol{\theta}, \boldsymbol{\Omega}, \lambda \sim N(\mathbf{a}(\boldsymbol{\theta}), (\lambda T)^{-1} [\boldsymbol{\Omega} \otimes \boldsymbol{\Gamma}_{\mathbf{xx}}(\boldsymbol{\theta})^{-1}])$$

What is λ ?

(1) parameter of prior

bigger λ means more confident in prior info.

If earlier we'd observed a sample of λT observations for which OLS estimate was $\mathbf{a}(\boldsymbol{\theta})$ and for which we'd had a diffuse prior, our posterior would have this form

$$\mathbf{a}|\boldsymbol{\theta}, \boldsymbol{\Omega}, \lambda \sim N(\mathbf{a}(\boldsymbol{\theta}), (\lambda T)^{-1} [\boldsymbol{\Omega} \otimes \boldsymbol{\Gamma}_{\mathbf{xx}}(\boldsymbol{\theta})^{-1}])$$

λ reflects our confidence in DSGE

$\lambda = 1$ DSGE counts just as much as current sample

$\lambda = 2$ DSGE counts twice as much as current sample

$\lambda = 0$ DSGE counts for nothing

If sample T is twice as big, I have twice as much confidence in the prior.

approximation?

useful rule of thumb?

What is λ ?

(2) hyperparameter

$$\mathbf{a} = \mathbf{a}(\boldsymbol{\theta}) + \mathbf{z}$$

$$\mathbf{z} \sim N(\mathbf{0}, (\lambda T)^{-1} [\boldsymbol{\Omega} \otimes \boldsymbol{\Gamma}_{\mathbf{xx}}(\boldsymbol{\theta})^{-1}])$$

Define $\hat{\theta}$ to be the QMLE:

$$\hat{\theta} = \arg \max \mathcal{L}(\theta)$$

$$\mathcal{L}(\theta) = -(T/2) \ln |\Omega(\theta)|$$

$$- (1/2) \sum_{t=1}^T \mathbf{q}_t' \Omega(\theta)^{-1} \mathbf{q}_t$$

$$\mathbf{q}_t = \mathbf{y}_t - \mathbf{A}(\theta) \mathbf{x}_t$$

$$\theta_0 = \text{plim } \hat{\theta}$$

$$\mathbf{A}_0 = \text{plim } \hat{\mathbf{A}}$$

prior: $\mathbf{a}|\boldsymbol{\theta}, \boldsymbol{\Omega}, \lambda \sim$

$$N(\mathbf{a}(\boldsymbol{\theta}), (\lambda T)^{-1} [\boldsymbol{\Omega}(\boldsymbol{\theta}) \otimes \boldsymbol{\Gamma}_{\mathbf{xx}}(\boldsymbol{\theta})^{-1}])$$

λ measures how far we expect \mathbf{a}_0 to be from $\mathbf{a}(\boldsymbol{\theta}_0)$ before seeing data

posterior distributions:

$$\boldsymbol{\theta}|\mathbf{Y}$$

$$\mathbf{a}|\mathbf{Y}$$

$$(\boldsymbol{\theta}, \mathbf{a})|\mathbf{Y}$$

$$E[\mathbf{a} - \mathbf{a}(\boldsymbol{\theta})][\mathbf{a} - \mathbf{a}(\boldsymbol{\theta})]'|\mathbf{Y}$$

could calculate ratio of determinant
of last magnitude to that of

$$E[\mathbf{a} - \mathbf{a}_0][\mathbf{a} - \mathbf{a}_0]'|\mathbf{Y}$$