Discussion of

What You Match Does Matter: The Effects of Data on DSGE Estimation

by Pablo Guerron-Quintana

Marc Giannoni
Columbia University, NBER and CEPR

Workshop on Methods and Applications for DSGE Models
Federal Reserve Bank of Cleveland
October 12, 2007
What does the paper do?

- DSGE Model (CEE, SW, DSSW, JP,…)
  
  Model: \( S_t = A(\theta) S_{t-1} + B(\theta) \varepsilon_t \)

  Selected endog. var.: \( F_t = C(\theta) S_t \)

  Observation equ.: \( X_t = F_t + \varepsilon_t \)

- Estimation:
  - Most recent papers use 7-9 observable series \( (X_t) \)
    - One for each concept: Y, C, I, L, W/P, \( \pi \), R
  - Here: systematically suppress one variable \( (X_{it}, F_{it}, \varepsilon_{it}) \) at a time and re-estimate

- Goal: Characterize impact of indicators \( X_t \) for estimation
  - on structural parameters \( \theta \) (median and dispersion)
  - on responses of model variables \( (S_t, F_t) \) to structural shocks \( \varepsilon_t \)
Paper’s findings

• Choice of observable series \( (X_t) \) plays an important role
  1. Absence of indicators of \( R, \pi, W/P \) affects importantly estimated persistence
     • Habit: \( 0.7 < h < 0.97 \)
     • Coeff on \( R(t-1) \) in Taylor rule: \( 0.06 < \rho_r < 0.87 \)

  2. Several parameters are robust to choice of variables
     • Steady-state values of \( \pi, L, \) growth rate, \( \beta \)

  3. Variance of parameter estimates affected
     • Excluding labor or investment greatly increases uncertainty of parameter estimates

  4. Estimates of shocks strongly affected
     • Related to Justiniano-Primiceri-Tambalotti (2006)

  5. IRFs strongly affected by choice of series
     • Without interest rate in estimation \( \Rightarrow \) strong deflation follows monetary expansion
My Comments

• Very nice paper
  – Useful systematic analysis of role of data series used in estimation

• Important message
  – Be careful not to omit relevant information in estimating model!

• Rest of my discussion
  – A few minor quibbles
  – Implications of results
  – Intuition for choice of number of series
  – Conclusion
Quibble 1: Some parameters are “robust to exclusion of observable variables”

• Guess which?

Table 1: Priors Densities for Structural Parameters

<table>
<thead>
<tr>
<th>$\sigma_m$</th>
<th>$\sigma^g_L$</th>
<th>$\sigma^g_K$</th>
<th>$\sigma^c_L$</th>
<th>$b$</th>
<th>$\xi_w$</th>
<th>$\xi_p$</th>
<th>$\gamma$</th>
<th>$\rho_R$</th>
<th>$\phi_\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IG [2.1]</td>
<td>IG [2.1]</td>
<td>IG [2.1]</td>
<td>IG [4.2]</td>
<td>B</td>
<td>B [0.6,0.1]</td>
<td>B [0.6,0.1]</td>
<td>B [1.0,1.1]</td>
<td>B [0.75,0.1]</td>
<td>B [1.60,0.3]</td>
</tr>
<tr>
<td>$\phi_y$</td>
<td>$\kappa$</td>
<td>$S_{U^c}$</td>
<td>$g_L$</td>
<td>$\pi$</td>
<td>$L$</td>
<td>$\zeta$</td>
<td>$\zeta_w$</td>
<td>$\kappa_a$</td>
<td>$\beta$</td>
</tr>
<tr>
<td>$G$ [0.12,0.1]</td>
<td>$N$ [3.1]</td>
<td>$N$ [1.3,0.3]</td>
<td>$N$ [1.01,0.003]</td>
<td>$N$ [1.01,0.002]</td>
<td>$N$ [5289.3]</td>
<td>$N$ [6.7,1]</td>
<td>$N$ [6.7,1]</td>
<td>$N$ [0.5,0.1]</td>
<td>$B$ [0.99,0.002]</td>
</tr>
<tr>
<td>$\rho_{gL}$</td>
<td>$\rho_{gK}$</td>
<td>$\rho_{U^c}$</td>
<td>$g_K$</td>
<td>$\sigma_{out}$</td>
<td>$\sigma_{cons}$</td>
<td>$\sigma_{invest}$</td>
<td>$\sigma_{labor}$</td>
<td>$\sigma_{wage}$</td>
<td>$\sigma_{inflat}$</td>
</tr>
<tr>
<td>$B$ [0.5,0.15]</td>
<td>$B$ [0.5,0.15]</td>
<td>$B$ [0.5,0.15]</td>
<td>$N$ [1.01,0.003]</td>
<td>$IG$</td>
<td>$IG$</td>
<td>$IG$</td>
<td>$IG$</td>
<td>$IG$</td>
<td>$IG$</td>
</tr>
</tbody>
</table>

Notes: $IG\sim$Inverse Gamma, $B\sim$Beta, $N\sim$Normal, $G\sim$Gamma

Mean and Standard Deviation in square brackets

• Very tight priors?
• Sounds more like an assumption than a result
Quibble 2 (smiS): Trend?

• Model includes 2 trends

• But data seems to have 3 trends (DSSW)

• Model surely misspecified

• Trend likely to play an important role in estimation

• Not clear a priori that provides reasonable estimates for fluctuations around trend
  – Possible solution: estimate by choosing frequency (Schorfheide et al.?)
Quibble 3: Stability over sample?

- Estimate model over 1954-2004

- But stability?
  - E.g., of policy rule around 1980?
    - Debate remains… but I find it difficult not to see a change (e.g. Boivin, 2005, JMCB)
  - Shock volatility?
    - E.g., Justiniano-Primiceri (2007)
Pablo’s proposal: 
Estimate with 7-8 series

• Intuition for using **more** variables
  – “Omission of relevant variables [i.e., observable indicators] leads to biased coefficients” (p. 3)
  – Yes! I fully agree

• Pablo’s proposal:
  – 1\textsuperscript{st} best: estimate model with measures of Y, C, I, L, W/P, π, R, Price of investment (but discontinued)
  – So 2\textsuperscript{nd} best: **estimate with remaining series**

• But…This is what SW and many others do!

• So are we done? (2\textsuperscript{nd} best achieved)

• Pablo seems to say “Yes”. His intuition for not using too many series:
  – “we know from regression analysis that adding more observables fictitiously improves model fitting at the cost of estimating more imprecise parameters” (p. 3)
  – No! Regression analogy does not apply here. Why?
Why might 7-8 series not be enough?

Inflation: Which series? CPI or PCE defl?
One proposal: DSGE model in data-rich environment (Boivin-Giannoni, 2006)

- **Model:**
  \[ S_t = A(\theta) S_{t-1} + B(\theta) \varepsilon_t \]

- **Selected endog. var.:**
  \[ F_t = C(\theta) S_t \]

- But generalize observation equation to:
  \[ X_t = \Lambda S_t + e_t \]
  - \( X_t \) = potentially large vector of data series [e.g. \( \Delta \) log(CPI), \( \Delta \) log(PCE defl.)]
  - \( S_t \) = latent state vector (satisfies restrictions imposed by model)
  - \( e_t \) = series-specific component

- **Note:** The more data \( X \), the more the model needs to explain
  - “adding more observables fictitiously improves model fitting at the cost of estimating more imprecise parameters” is **NOT TRUE**
Estimated Inflation:
Properties vary a lot whether 1 or more inflation series are used

Quarterly inflation (demeaned)

A. Meas. Error (as in Sargent, 1989)

B. Multiple infl. indicators

C. Large data set
Large data set
Estimated Inflation:
More precisely estimated with large data set

Meas. Error (as in Sargent, 1989):
Imprecisely estimated; some parameters not identified

Large data set:
Inflation PRECISELY estimated
DSGE model in data-rich environment

- Other results (BG, 2006):
  - Sources of BC fluctuations depend on data set considered
    - E.g., if $\Delta \log(\text{GDP defl.}) = \pi$
      - $\pi$ largely explained by markup shocks
    - If $\pi$ is estimated from several indicators
      - Markup shocks less important
  - Parameters estimates affected
Conclusion

• Very nice paper
• Important message
  – Be careful not to omit relevant information in estimating model!

But as Thompson and Thomson say:

“I fully agree, but to be precise… “
Conclusion

• We can (and should) go further!
  – Use multiple indicators (if available) for latent concepts
  – No reason to omit available data if it is informative about model concepts
  – In the same way that:
    • data on inflation is useful to the estimation of inflation persistence,
    • data on interest rates is relevant to the estimation of interest rate rule
    • data on wages is relevant to estimation of wage persistence/stickiness

  ... a large number of macro indicators should be useful for estimation of productivity, beliefs, capital stock....