Optimal Monetary Policy in a Data-Rich Environment

Jean Boivin  
HEC Montréal,  
CIRANO, CIRPÉE and NBER

Marc Giannoni  
Columbia University,  
NBER and CEPR

“Forecasting Short-term Economic Developments…”  
Bank of Canada  
October 25-26, 2007
Importance of large data sets: Evidence from factor models

- **Forecasting**
  

- **Monetary policy**
  
  [Bernanke and Boivin (2003), Giannone, Reichlin and Sala (2004)]

- **VAR**
  
  [Bernanke, Boivin and Eliasz (2005), Forni, Giannone, Lippi, Reichlin (2004)]
Factor Augmented Vector Autoregression: FAVAR

Bernanke, Boivin and Eliasz (2005)

Observation equation:

\[ X_t = \Lambda C_t + e_t \]

Transition equation (VAR):

\[ C_t = \Phi(L)C_{t-1} + v_t \]

\( R \) — set of observable series (here Fed funds rate)

\( F \) — set of unobservable factors ("economic activity," "inflation," ...)

\( X \) — large panel of informational series

\( e \) — series-specific component (potentially serially and weakly cross correlated)
Estimated responses to an identified monetary policy shock

- Federal Funds Rate
- Industrial Production
- Price Level (log): CPI

- Blue line: Baseline FAVAR (5 factors)
- Green dashed line: VAR [Indus. prod., CPI, FFR]
- Red line: VAR & 1 factor
Why are a large set of macro indicators useful?

Existing evidence on use of based on largely non-structural models. Limits our ability to:

- Determine why large data sets are useful
- Determine sources of fluctuations
- Perform counterfactual experiments
- Analyze optimal policy
Estimated DSGE models

- Important developments

- Now increasingly taken seriously as empirical models

- Promising empirical success

  [Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2004)]

- Estimated based on a handful of data series
  \[\rightarrow\] at odds with fact that CB and financial market participants monitor large number of data series!
Goal of the research agenda

“The more specific and data-rich the model, the more effective it will be” (Greenspan’s memoirs, 2007)

- To explore role of large data sets for estimated DSGE models
- By product: Provide economic interpretation of latent factors
- Optimal monetary policy in a data-rich environment
Preview of the main results

- More precise estimation of the state of the economy

- Improvements in “forecasting” with additional information

- Different conclusions about structure of economy and sources of business cycles
  - Different propagation mechanism (e.g. less habit formation and inflation indexing)
  - Fewer and smaller structural shocks

- Information from large data set might matter for optimal policy
Why more data in a DSGE context?
Why more data in a DSGE context?

- No scope if:
  - Model well specified
  - Theoretical concepts directly observed by agents and econometrician
Why more data in a DSGE context?

- No scope if:
  - Model well specified
  - Theoretical concepts directly observed by agents and econometrician

- Empirical evidence: at least one assumption violated
Why more data in a DSGE context?

- No scope if:
  - Model well specified
  - Theoretical concepts directly observed by agents and econometrician

- Empirical evidence: at least one assumption violated

- We assume theoretical concepts partially observed by econometrician
  - Employment: Discrepancies between household and payroll surveys
  - Inflation: GDP deflator, CPI
  - Productivity shock: oil prices, commodity prices

- If indeed we are missing information in DSGE estimation: all parameter estimates potentially distorted!
Outline of presentation

- Data-rich environment
  - A simple example: RBC model
  - General framework
  - Estimation

  - Results

- Optimal monetary policy

- Conclusion
Why more information? A simple RBC model

Households maximize lifetime utility

\[ E_0 \sum_{t=0}^{\infty} \beta^t \left[ \log (c_t) + \nu \log (1 - l_t) \right], \quad 0 < \beta < 1, \ \nu > 0 \]

subject to

\[ y_t = e^{a_t} k_t^{1-\alpha} l_t^{\alpha}, \quad 0 < \alpha < 1 \]
\[ y_t = c_t + k_{t+1} - (1 - \delta) k_t, \]
\[ a_t = \rho a_{t-1} + \varepsilon_t, \quad 0 \leq \rho < 1, \ \varepsilon_t \sim N(0, \sigma). \]
RBC example continued...

- Linearized solution has the form:

\[
\begin{align*}
\hat{y}_t &= d_1 \hat{k}_t + d_2 a_t \\
\hat{c}_t &= d_3 \hat{k}_t + d_4 a_t \\
\hat{l}_t &= d_5 \hat{k}_t + d_6 a_t \\
\hat{k}_t &= g_1 \hat{k}_{t-1} + g_2 a_{t-1} \\
a_t &= \rho a_{t-1} + \varepsilon_t
\end{align*}
\]

\[
\begin{align*}
z_t &= DS_t, \\
z_t &= [\hat{y}_t, \hat{c}_t, \hat{l}_t]' \\
S_t &= GS_{t-1} + H\varepsilon_t, \\
S_t &= [\hat{k}_t, a_t]'
\end{align*}
\]

where \(D, G\) and \(H\) are functions of model parameters

- Suppose we estimate the model on the basis of only one variable (no stochastic singularity):

\[
F_t = \hat{y}_t = [d_1 \ d_2] S_t
\]
**RBC example: How to link model and data?**

- One indicator, no measurement error

\[ X_t = \hat{y}_t = d_1 \hat{k}_t + d_2 a_t \]

- e.g. \( X_t = \) real GDP

- No scope for judgment or soft data

- One indicator, measurement error (Sargent, 1989):

\[ X_t = \hat{y}_t + e_t = d_1 \hat{k}_t + d_2 a_t + e_t \]

- 1 shock and 1 measurement error: identification from dynamics

- Identification problems?
RBC example (cont.): Proposed solution

- Multiple indicators with **known** relationships to a theoretical concepts

\[ \begin{align*}
X_t &= \begin{bmatrix} \text{real GDP} \\ \text{real NI} \end{bmatrix} = \begin{bmatrix} 1 \\ \lambda_{NI} \end{bmatrix} \hat{y}_t + e_t \\
&= \begin{bmatrix} d_1 \\ \lambda_{NI}d_1 \\ d_2 \\ \lambda_{NI}d_2 \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ a_t \end{bmatrix} + e_t
\end{align*} \]

- Helps disentangle meas. error from structural shocks

- Multiple indicators with **unknown** link

- E.g., soft data

\[ \begin{align*}
X_t &= \begin{bmatrix} \text{real GDP} \\ \text{soft data} \end{bmatrix} = \begin{bmatrix} d_1 & d_2 \\ \lambda_3 & \lambda_4 \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ a_t \end{bmatrix} + e_t
\end{align*} \]
**Benefits of exploiting more information: Intuition**

- Measurement error identifiable from cross-section of indicators

  Example: $x_{it} = f_t + e_{it}$, $i = 1, ..., n_X$

  - If $n_X = 1$, and both $f_t$ and $e_{it}$ are i.i.d. $\implies$ Not identified
  - If $n_X = 1$, $f_t$ is AR(1) and $e_{it}$ is i.i.d. $\implies$ Identified (from dynamics)
  - If $n_X > 1$, and both $f_t$ and $e_{it}$ are i.i.d. $\implies$ Identified (from cross-section)

- Permits the identification of more structural shocks

- Don’t have to take a stand *a priori* on the relative importance of measurement errors vs structural shocks

- More efficient (consistent) estimate of the latent factors

  - $\text{var}(\hat{f}_t)$ is of order $1/n_X$ [Stock Watson (2002), Forni et al. (2000)]
Empirical model: Summary

- Transition equation:
  \[ S_t = GS_{t-1} + H\varepsilon_t \]

- Observation equation:
  \[ X_t = \Lambda S_t + e_t \]
  where
  \[ X_t \equiv \begin{bmatrix} X_{F,t} \\ X_{S,t} \end{bmatrix}, \quad e_t \equiv \begin{bmatrix} e_{F,t} \\ e_{S,t} \end{bmatrix}, \quad \Lambda \equiv \begin{bmatrix} \Lambda_F \Phi \\ \Lambda_S \end{bmatrix}. \]

- Comments:
  - Related to non-structural factor models, but we impose DSGE model on transition equation of latent factors
  - Factors have economic interpretation: state variables
  - Interpret info. in data set through lenses of DSGE model
  - Can do counterfactual experiments, study optimal policy
Application: Smets and Wouters (2004) [i.e., CEE (2005) with shocks]

- State-of-the-art DSGE model:
  - Popular as fits apparently well, good for forecasting
  - Many frictions, many shocks

- Households
  - Consume aggregate of all goods, habit formation (external)
  - Supply specialized labor on monopolistically competitive labor mkt
  - Rent capital services to firms
  - Decide how much capital to accumulate

- Firms:
  - Choose labor and capital inputs
  - Supply differentiated goods on monopolistically competitive goods mkt

- Prices and wages reoptimized at random intervals (Calvo)
  - If not reoptimized: indexed to past inflation and CB’s inflation target
Smets and Wouters (2004): Model solution

- 7 variables of interest: \( F_t = [i_t, Y_t, C_t, I_t, \pi_t, w_t, L_t]' \)

- 9 shocks: \( s_t = [\varepsilon_t^a, \varepsilon_t^b, \varepsilon_t^G, \varepsilon_t^L, \varepsilon_t^I, \eta_t^Q, \eta_t^p, \eta_t^w, \eta_t^i]' \)

- State vector

\[
S_t = [i_{t-1}, Y_{t-1}, C_{t-1}, I_{t-1}, \pi_{t-1}, w_{t-1}, K_{t-1}, \varepsilon_{t-1}^I, \pi_{t-1}, s_t]'
\]

- State-space representation:
  - Transition equation

\[
S_t = G S_{t-1} + H \varepsilon_t
\]

  - Observation equation

\[
X_t = \Lambda S_t + e_t
\]
Estimation method

- Difficult problem to estimate (large dimension)

- Standard methods not successful (e.g. MLE)

- MCMC methods:
  - Empirical approximation of the posterior distribution. Does not rely on gradient method
  - Draw iteratively from conditional distributions (solves the high-dimensionality problem)
  - Priors can help make the estimation better behaved
• Specifications of observation equation: \( X_t = \Lambda S_t + e_t \)

• **Case SW**: Standard estimation (as in Smets and Wouters)

\[
X_{F1,t} = F_t = \Phi S_t
\]

where

\[
X_{F1,t} = [\text{Fed funds, GDP, cons., invest., } \%\Delta GDP \text{ defl, real wage, hours worked}]'
\]

• **Case A** = Case SW + Measurement error (as in Sargent, 1989):

\[
X_{F1,t} = F_t + e_t = \Phi S_t + e_t
\]

Restrictions of model used to estimate latent variables in \( F_t \) (identification problems?)
Specifications of observation equation (cont.)

- **Case B** = Case A + 7 new indicators of $F_t$ (14 series in total)

\[ X_{F,t} = \Lambda_F F_t + e_t = \Lambda_F \Phi S_t + e_t \]

\[ X_{F,t} = \left[ X'_{F1,t}, X'_{2,t} \right]' \]

\[ X_{2,t} = \left[ \text{cons. excl. food & energy, priv. invest.,} \right. \]
\[ \text{CPI, core CPI, PCE defl, empl. (HH and est. surveys)} \left. \right] ' \]

- E.g. for inflation: use GDP defl., PCE defl., CPI

- **Case C** = Same as case B, but with unrestricted loading matrix larger data set (99 series)

\[ X_{F1,t} = F_t + e_{F1,t} = \Phi S_t + e_{F1,t} \]

\[ X_{2,t} = \Lambda_S S_t + e_{S,t} \]

\[ \left\{ X_t = \Lambda S_t + e_t \right\} \]
Evidence of “measurement errors”

Distribution of correlations between latent concepts and reference indicators
Empirical results: Estimated latent variables

- Interest rate: $i$
- Output: $Y$
- Consumption: $C$
- Investment: $I$
- Real wage: $w$
- Employment: $L$

Data A B C
Empirical results: Estimated inflation
Estimated Inflation: Median, 5th and 95th percentiles
More information leads to more precise estimates of the latent variables

<table>
<thead>
<tr>
<th>Concept</th>
<th>Case A st. dev.</th>
<th>Case B Relative to case A</th>
<th>Case C Relative to case A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate $R_t$</td>
<td>0.000</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Output $Y_t$</td>
<td>0.342</td>
<td>0.93</td>
<td>1.01</td>
</tr>
<tr>
<td>Consumption $C_t$</td>
<td>0.450</td>
<td>0.93</td>
<td>1.01</td>
</tr>
<tr>
<td>Investment $I_t$</td>
<td>0.908</td>
<td>0.94</td>
<td>0.89</td>
</tr>
<tr>
<td>Inflation (annualized) $\pi_t$</td>
<td>0.500</td>
<td>0.91</td>
<td>0.65</td>
</tr>
<tr>
<td>Real wage $w_t$</td>
<td>0.478</td>
<td>1.04</td>
<td>1.06</td>
</tr>
<tr>
<td>Hours worked $L_t$</td>
<td>0.311</td>
<td>0.76</td>
<td>0.97</td>
</tr>
</tbody>
</table>
"Forecasting" performance: One-step ahead RMSE’s

<table>
<thead>
<tr>
<th>Primary indicator</th>
<th>Case A RMSE</th>
<th>Case B Relative to case A</th>
<th>Case C Relative to case A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fed funds rate</td>
<td>0.52</td>
<td>1.08</td>
<td>1.12</td>
</tr>
<tr>
<td>Real GDP</td>
<td>0.55</td>
<td>1.00</td>
<td>1.02</td>
</tr>
<tr>
<td>Real Consumption</td>
<td>0.59</td>
<td>0.93</td>
<td>0.97</td>
</tr>
<tr>
<td>Real Investment</td>
<td>1.64</td>
<td>0.97</td>
<td>0.88</td>
</tr>
<tr>
<td>GDP defl. inflation</td>
<td>0.20</td>
<td>0.95</td>
<td>0.90</td>
</tr>
<tr>
<td>Real wage</td>
<td>0.75</td>
<td>1.03</td>
<td>0.96</td>
</tr>
<tr>
<td>Hours worked</td>
<td>0.49</td>
<td>1.02</td>
<td>1.04</td>
</tr>
<tr>
<td>Overall</td>
<td>-9.26</td>
<td>0.98</td>
<td>0.98</td>
</tr>
</tbody>
</table>
Benefit of adding more information

- More precise estimates of latent variables, in particular inflation

- Better “forecasts” of 7 reference series
Correlation between observable indicators and corresponding latent concepts

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fed funds rate (120)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Real GDP (1)</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>Real Consumption (49)</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Real fixed Investment (74)</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>GDP defl. inflation (145)</td>
<td>0.71</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>Real wage (18)</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Hours worked (23)</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>PCE ex. food and Energy (71)</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>Gross Real Investment (73)</td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>PCE deflator (146)</td>
<td>0.68</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>core-CPI (208)</td>
<td>0.52</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td>CPI (215)</td>
<td>0.53</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Employment HH Survey (28)</td>
<td>0.89</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>Payroll Employment (36)</td>
<td>0.81</td>
<td>0.85</td>
<td>0.85</td>
</tr>
</tbody>
</table>
## Estimated structural parameters

<table>
<thead>
<tr>
<th></th>
<th>Prior Distribution</th>
<th>SW</th>
<th>Case A</th>
<th>Case B</th>
<th>Case C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td>Mean</td>
<td>St.Err.</td>
<td>Type</td>
<td>Mean</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>Normal</td>
<td>4</td>
<td>1.5</td>
<td></td>
<td>5.36</td>
</tr>
<tr>
<td>$h$</td>
<td>Beta</td>
<td>0.7</td>
<td>0.1</td>
<td></td>
<td>0.71</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Normal</td>
<td>1.25</td>
<td>0.125</td>
<td></td>
<td>1.42</td>
</tr>
<tr>
<td>$1/\psi$</td>
<td>Normal</td>
<td>0.2</td>
<td>0.075</td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>$\gamma_\omega$</td>
<td>Beta</td>
<td>0.5</td>
<td>0.15</td>
<td></td>
<td>0.39</td>
</tr>
<tr>
<td>$\gamma_p$</td>
<td>Beta</td>
<td>0.5</td>
<td>0.15</td>
<td></td>
<td>0.66</td>
</tr>
<tr>
<td>$r_{\pi0}$</td>
<td>Normal</td>
<td>1.8</td>
<td>0.1</td>
<td></td>
<td>1.78</td>
</tr>
<tr>
<td>$r_{\pi1}$</td>
<td>Normal</td>
<td>-0.3</td>
<td>0.1</td>
<td></td>
<td>-0.22</td>
</tr>
<tr>
<td><strong>Implied parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pseudo EIS: $\frac{1-h}{(1+h)\sigma_c}$</td>
<td></td>
<td>0.110</td>
<td>0.099</td>
<td>0.167</td>
<td>0.204</td>
</tr>
<tr>
<td>slope of PC: $\frac{(1-\beta\xi_p)(1-\xi_p)}{(1+\beta\gamma_p)\xi_p}$</td>
<td></td>
<td>0.011</td>
<td>0.007</td>
<td>0.012</td>
<td>0.018</td>
</tr>
</tbody>
</table>
Estimated time series of capital and shocks

- $K$: Capital stock
- $\epsilon^a$: Productivity shock
- $\epsilon^b$: Preference shock
- $\epsilon^G$: Gov. expenditure shock
- $\epsilon^L$: Labor supply shock
- $\epsilon^I$: Invest.-specific shock

SW A B C
Estimated time series of shocks

$\eta^Q$: Equity premium shock

$\eta^P$: Price markup shock

$\eta^\omega$: Wage markup shock

$\eta^i$: Monetary policy shock

SW A B C
Variance decompositions
Findings

Adding more information leads to:

- More precise estimates of the state of the economy (inflation)

- “Forecasting” performance: improvements

- Different conclusions about the nature of propagation and sources of business cycle fluctuations

To be investigated further...

- What matters: estimation or filtering?

- Do info from large data set matter for welfare?
Optimal Monetary Policy in a Data-Rich Environment


- GW (2002): General characterization of optimal target criterion

\[ a(L)i_t + B(L)E_t \left[ C \left( L^{-1} \right) (\tau_t - \tau^*_t) \right] = 0 \]

Desirable properties: determinacy, robustness to shock processes...

- Here: Use model and large data set to improve forecasts for implementation of policy
Implementation

- Calibrate standard DGSE model (Giannoni Woodford (2003)) and assume structural parameters are known (no estimation, just filtering)

- Different cases: Theoretical variables are unobserved by:
  - By central bank (asymmetric info case)
  - Both central bank and agents (symmetric info case)

- Investigate:
  - Optimal monetary policy
  - Welfare implication of a central bank that does not account for large information set
Preliminary results

(based on shortcuts and Giannoni and Woodford, 2003)

- Assume the true variables driving the economy are as estimated under case C

- Consider two cases:
  - Data-Rich CB: Central bank implements policy on the basis of the “true” inflation (case C)
  - Data-Poor CB: Central bank ignores data-rich environment and respond instead to actual data (GDP deflator)

- Comparing loss functions: 23% higher for Data-Poor CB

<table>
<thead>
<tr>
<th></th>
<th>var($\pi_t$)</th>
<th>var($Y_t$)</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case FI</td>
<td>2.0028</td>
<td>11.6477</td>
<td>3.6402</td>
</tr>
<tr>
<td>Case Al</td>
<td>2.6206</td>
<td>5.8146</td>
<td>4.4689</td>
</tr>
</tbody>
</table>
Avenues for future research

- Real time application with mixed frequencies