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

  [Stock and Watson (1999, 2002), Forni, Hallin, Lippi, Reichlin (2000)]

- 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):

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

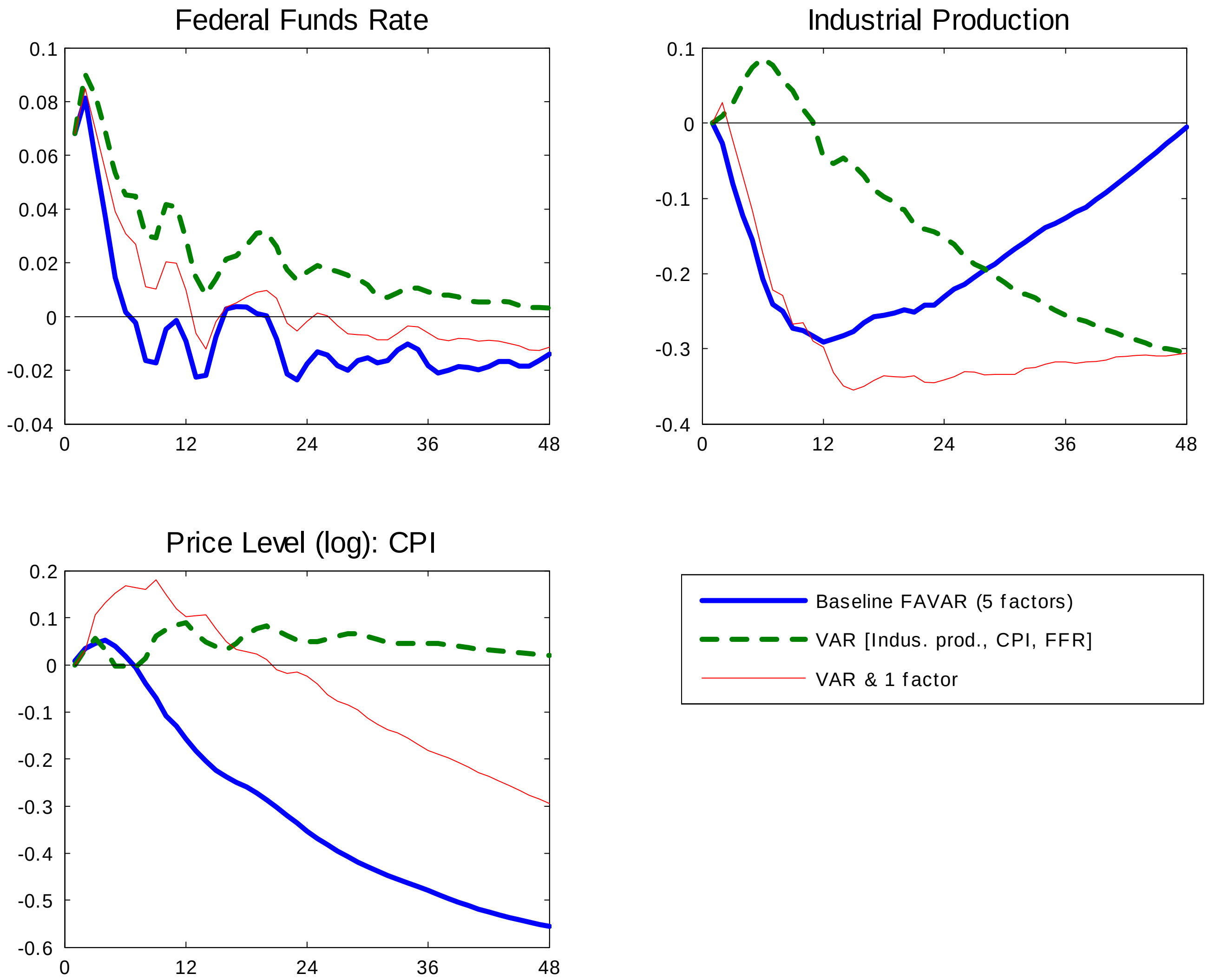
R — set of observable series (here Fed funds rate)

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



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
  [Altug (1989), McGrattan (1994), Leeper and Sims (1994), Rotemberg and Woodford (1997), Ireland (1997, 2001), Kim (2000), Schorfheide (2000), Christiano, Eichenbaum and Evans (2005), Amato and Laubach (2003), Smets and Wouters (2003, 2004), Altig, Christiano, Eichenbaum and Linde (2003), Rabanal and Rubio-Ramírez (2003), Julliard, Karam, Laxton and Pesenti (2004), LOWW (2005), Justiniano and Primiceri (2006), ...]

- 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
  $\Longrightarrow$ 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
- Application: Smets and Wouters (2004) model
  - Results
- Optimal monetary policy
- Conclusion

Why more information? A simple RBC model

Households maximize lifetime utility

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

subject to

$$\begin{aligned} y_t &= e^{a_t} k_t^{1-\alpha} l_t^{\alpha}, \qquad 0 < \alpha < 1 \\ y_t &= c_t + k_{t+1} - (1-\delta) k_t, \\ a_t &= \rho a_{t-1} + \varepsilon_t, \qquad 0 \le \rho < 1,\ \varepsilon_t \sim N(0, \sigma). \end{aligned}$$

RBC example continued...

- Linearized solution has the form:

$$\left.\begin{aligned}\hat{y}_t &= d_1\hat{k}_t + d_2 a_t\\ \hat{c}_t &= d_3\hat{k}_t + d_4 a_t\\ \hat{l}_t &= d_5\hat{k}_t + d_6 a_t\end{aligned}\right\}\quad z_t = DS_t, \qquad z_t = \left[\hat{y}_t,\ \hat{c}_t,\ \hat{l}_t\right]'$$

$$\left.\begin{aligned}\hat{k}_t &= g_1\hat{k}_{t-1} + g_2 a_{t-1}\\ a_t &= \rho a_{t-1} + \varepsilon_t\end{aligned}\right\}\quad S_t = GS_{t-1} + H\varepsilon_t, \qquad S_t = \left[\hat{k}_t, a_t\right]'$$

  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):

$$F_t = \hat{y}_t = \begin{bmatrix} d_1 & d_2 \end{bmatrix} S_t$$

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

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

  - Helps disentangle meas. error from structural shocks

- Multiple indicators with **unknown** link

- E.g., soft data

$$X_t = \begin{bmatrix} \text{real GDP} \\ \text{soft data} \end{bmatrix} = \begin{bmatrix} d_1 & d_2 \\ \lambda_3 & \lambda_4 \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ a_t \end{bmatrix} + e_t$$

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

  - $var(\hat{f}_t)$ 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

$$X_t \equiv \begin{bmatrix} X_{F,t} \\ X_{S,t} \end{bmatrix}, \qquad e_t \equiv \begin{bmatrix} e_{F,t} \\ e_{S,t} \end{bmatrix}, \qquad \Lambda \equiv \begin{bmatrix} \Lambda_F \Phi \\ \Lambda_S \end{bmatrix}.$$

- 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: $s_t = \left[\varepsilon_t^a, \varepsilon_t^b, \varepsilon_t^G, \varepsilon_t^L, \varepsilon_t^I, \eta_t^Q, \eta_t^p, \eta_t^w, \eta_t^i\right]'$

- State vector

$$S_t = \left[i_{t-1}, Y_{t-1}, C_{t-1}, I_{t-1}, \pi_{t-1}, w_{t-1}, K_{t-1}, \varepsilon_{t-1}^I, \bar{\pi}_{t-1}, s_t'\right]'$$

- State-space representation:
  - Transition equation

$$S_t = GS_{t-1} + H\varepsilon_t$$

  - 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)

$$X_{F\mathbf{1},t} = F_t = \Phi S_t$$

where

$$X_{F\mathbf{1},t} = [\text{Fed funds, GDP, cons., invest., }\%\Delta GDP\text{ defl, real wage, hours worked}]'$$

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

$$X_{F\mathbf{1},t} = F_t + e_t = \Phi S_t + e_t$$

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)

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

$$\begin{aligned} X_{F,t} &= \left[X'_{F1,t}, X'_{2,t}\right]' \\ X_{2,t} &= \left[\begin{array}{c} \text{cons. excl. food \& energy, priv. invest.,} \\ \text{CPI, core CPI, PCE defl, empl. (HH and est. surveys)} \end{array}\right]' \end{aligned}$$

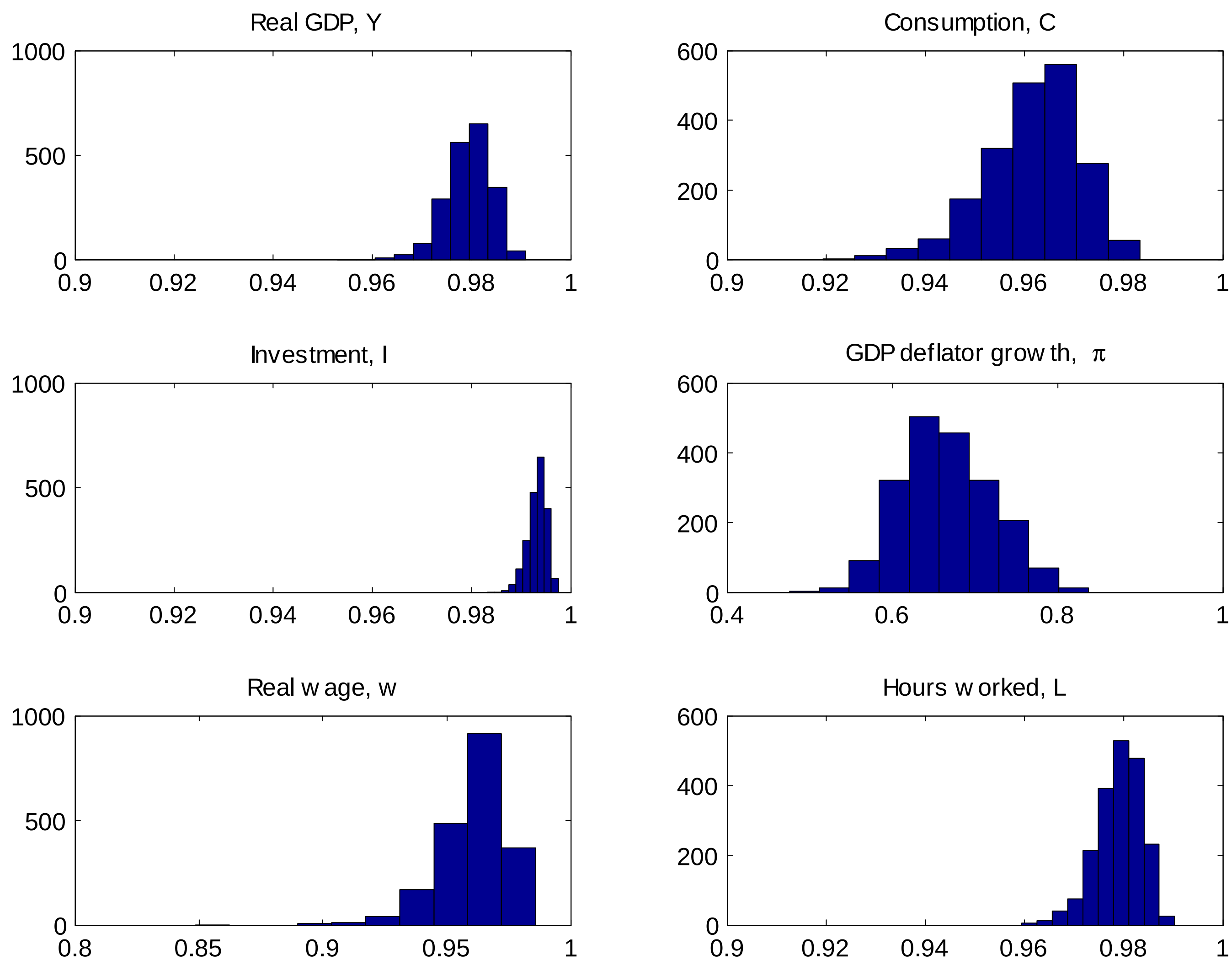
  – E.g. for inflation: use GDP defl., PCE defl., CPI

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

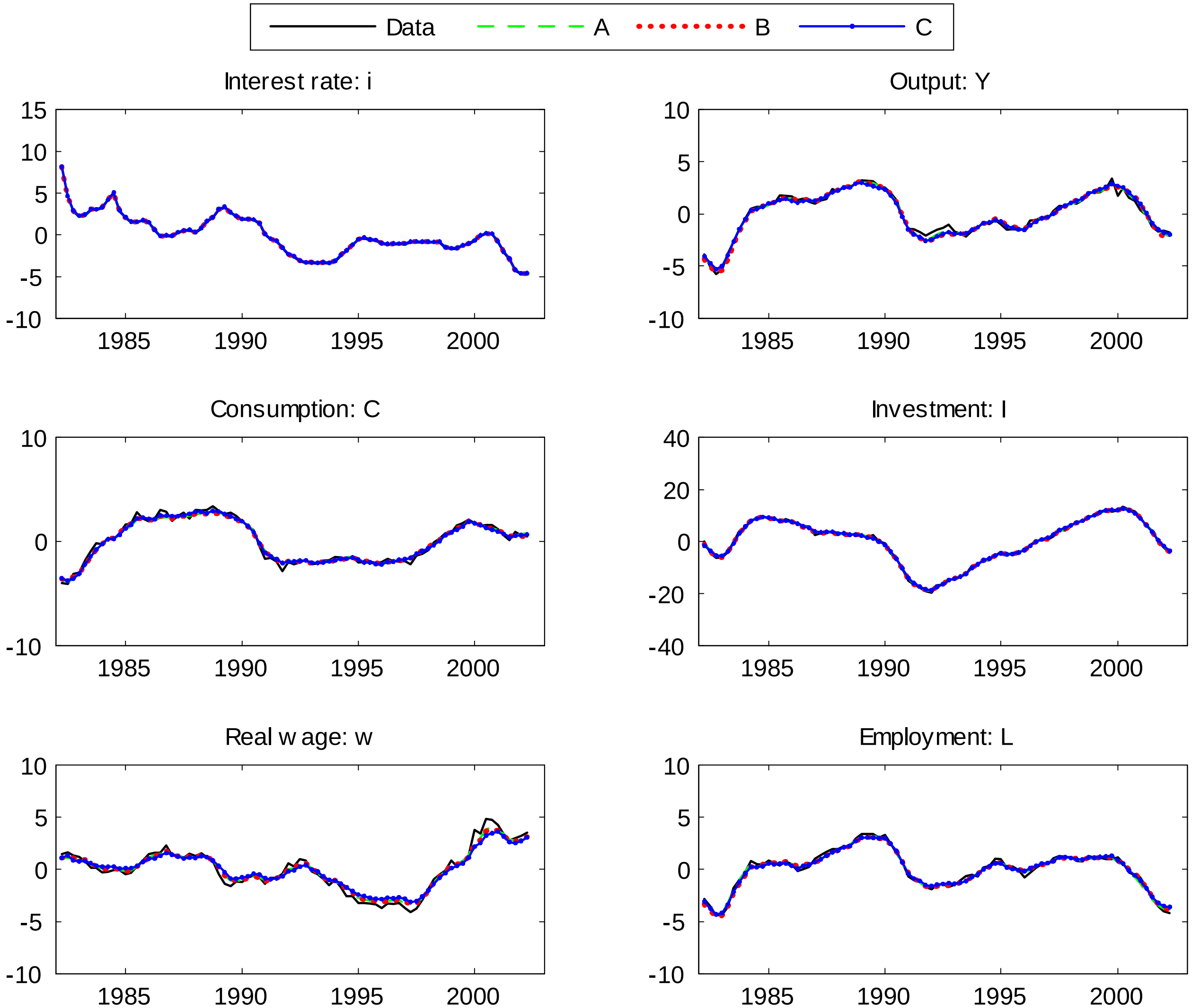
$$\left.\begin{array}{l} X_{F1,t} = F_t + e_{F1,t} = \Phi S_t + e_{F1,t} \\ X_{2,t} = \Lambda_S S_t + e_{S,t} \end{array}\right\} \Longleftrightarrow X_t = \Lambda S_t + e_t$$

Evidence of “measurement errors”

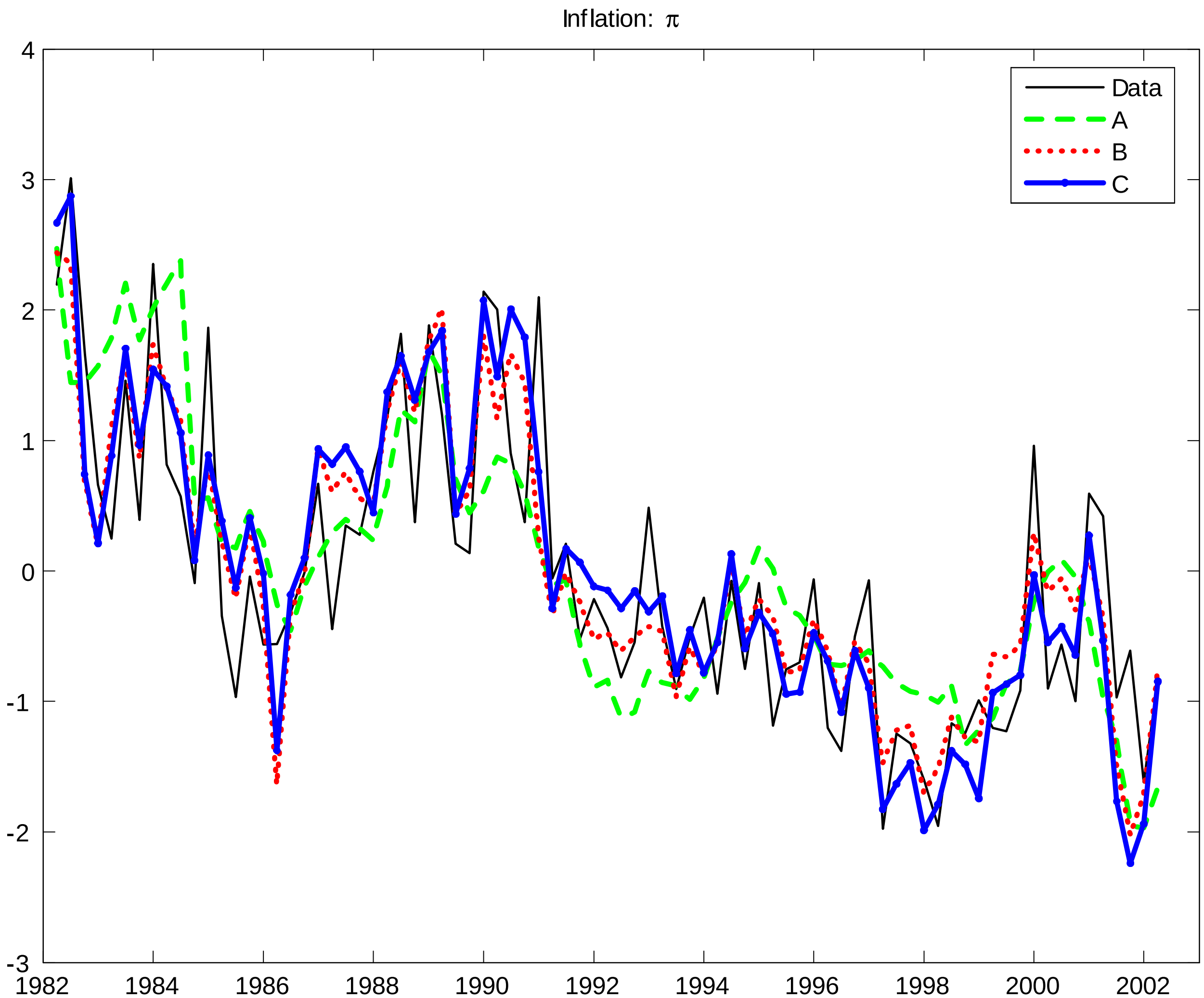
Distribution of correlations between latent concepts and reference indicators



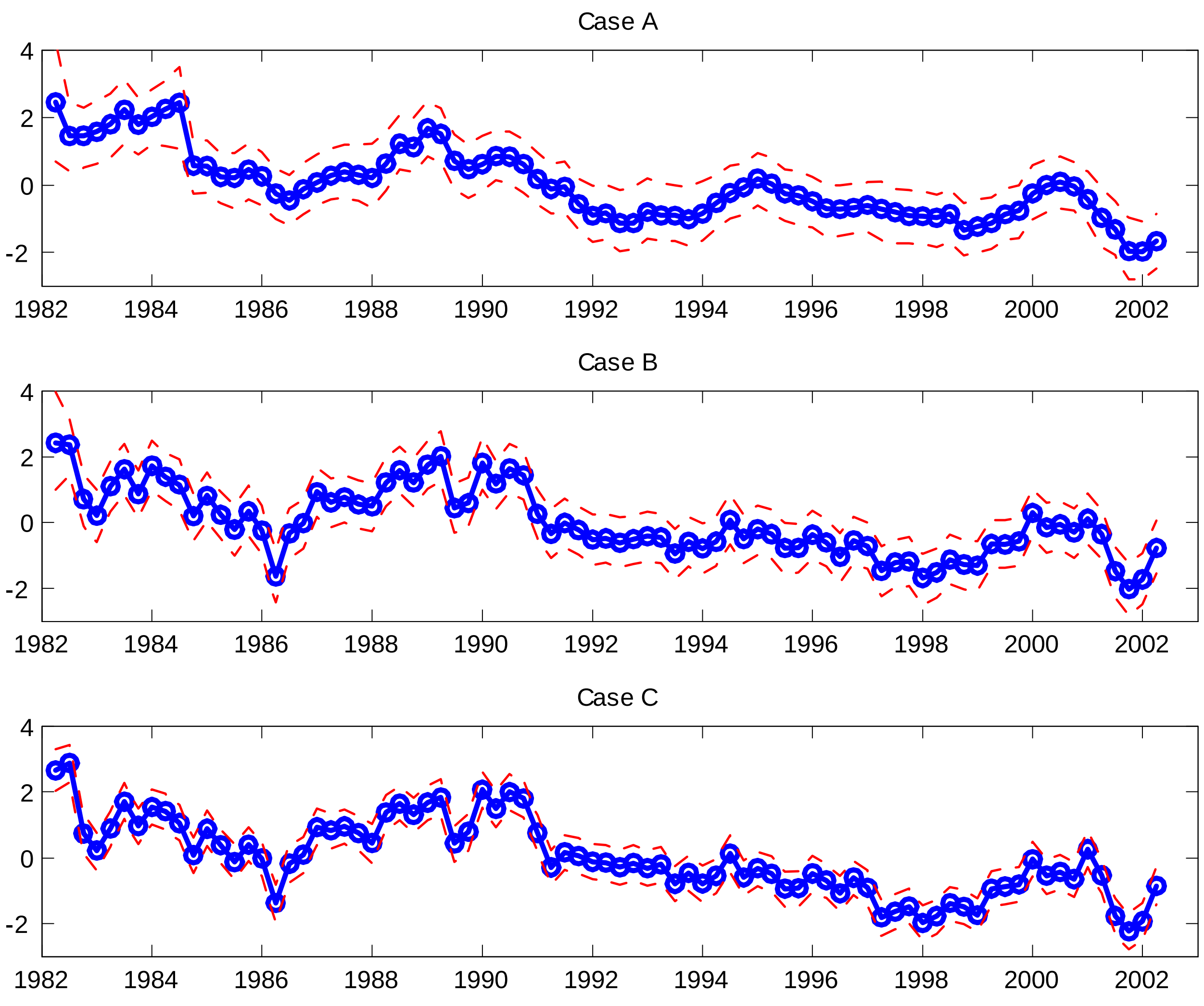
Empirical results: Estimated latent variables



Empirical results: Estimated inflation



Estimated Inflation: Median, 5th and 95th percentiles



More information leads to more precise estimates of the latent variables

Concept		Case A st. dev.	Case B	Case C
			Relative to case A	
Interest rate	R_t	0.000	—	—
Output	Y_t	0.342	0.93	1.01
Consumption	C_t	0.450	0.93	1.01
Investment	I_t	0.908	0.94	0.89
Inflation (annualized)	π_t	0.500	0.91	0.65
Real wage	w_t	0.478	1.04	1.06
Hours worked	L_t	0.311	0.76	0.97

“Forecasting” performance: One-step ahead RMSE’s

Primary indicator	Case A RMSE	Case B	Case C
		Relative to case A	
Fed funds rate	0.52	1.08	1.12
Real GDP	0.55	1.00	1.02
Real Consumption	0.59	0.93	0.97
Real Investment	1.64	0.97	0.88
GDP defl. inflation	0.20	0.95	0.90
Real wage	0.75	1.03	0.96
Hours worked	0.49	1.02	1.04
Overall	-9.26	0.98	0.98

Benefit of adding more information

- More precise estimates of latent variables, in particular inflation

- Better “forecasts” of 7 reference series

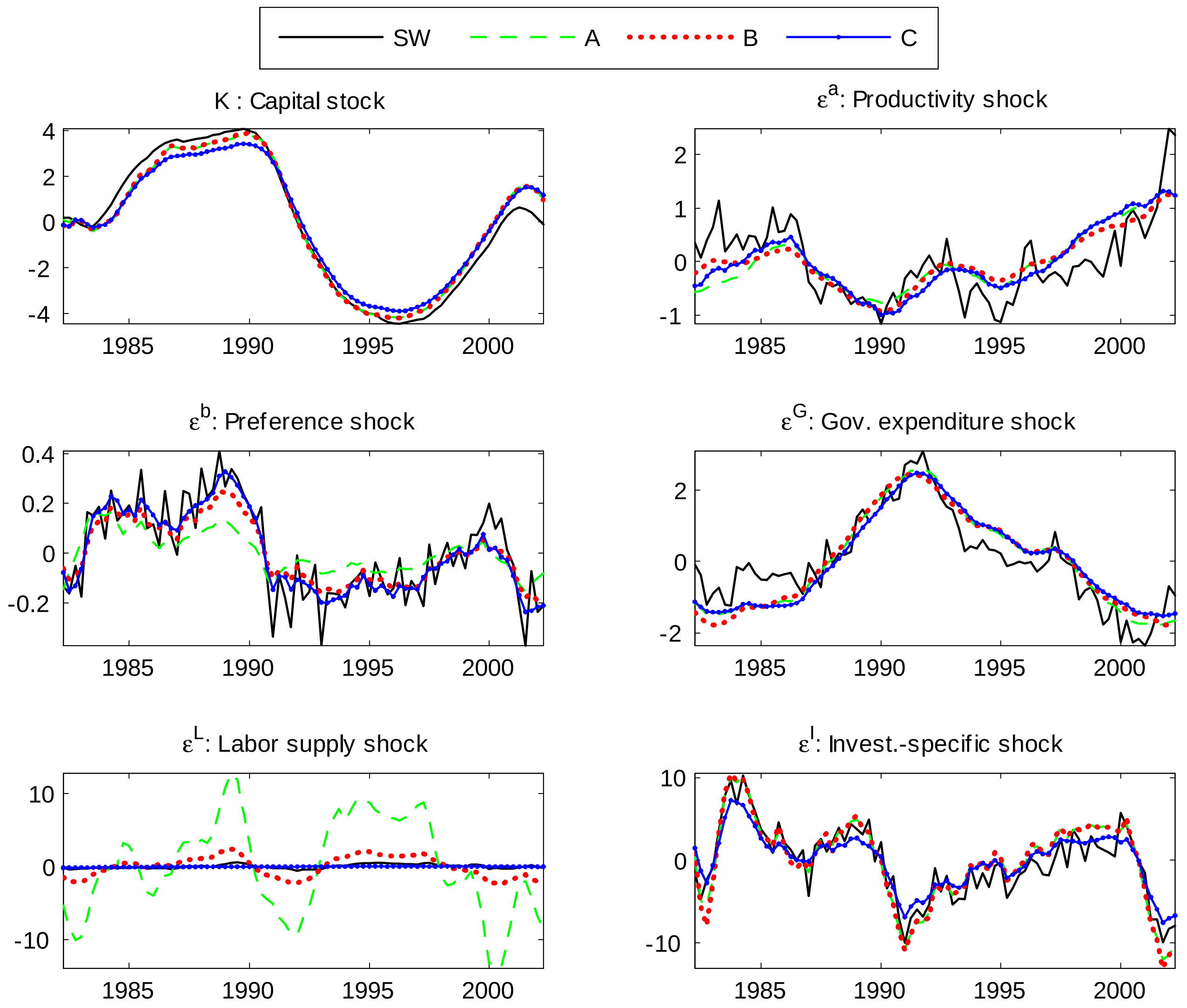
Correlation between observable indicators and corresponding latent concepts

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

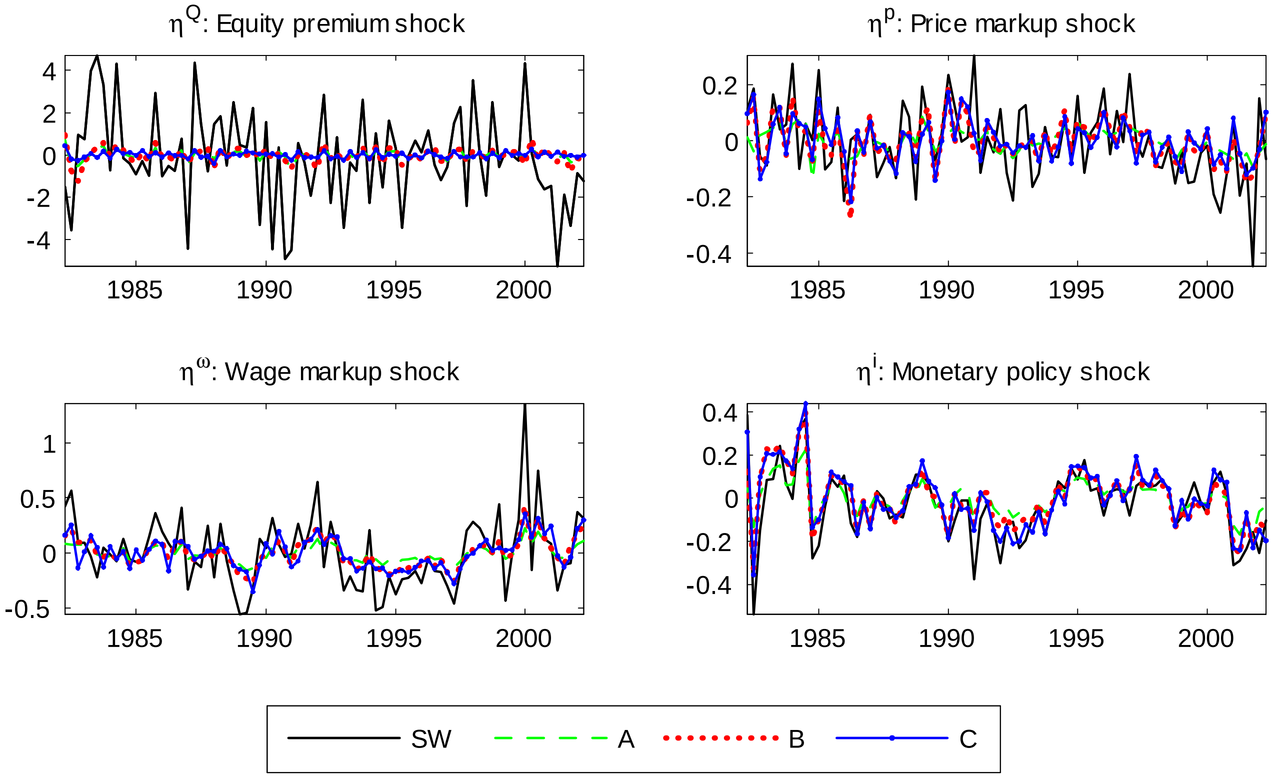
Estimated structural parameters

	Prior Distribution			SW	Case A	Case B	Case C
	Type	Mean	St.Err.				
φ	Normal	4	1.5	5.36	5.88	6.17	3.81
				(0.88)	(1.11)	(1.13)	(1.04)
h	Beta	0.7	0.1	0.71	0.75	0.54	0.50
				(0.07)	(0.07)	(0.27)	(0.27)
ϕ	Normal	1.25	0.125	1.42	1.24	1.37	1.26
				(0.08)	(0.07)	(0.07)	(0.07)
$1/\psi$	Normal	0.2	0.075	0.32	0.27	0.26	0.27
				(0.06)	(0.06)	(0.06)	(0.06)
γ_ω	Beta	0.5	0.15	0.39	0.45	0.43	0.48
				(0.12)	(0.14)	(0.14)	(0.14)
γ_p	Beta	0.5	0.15	0.66	0.72	0.50	0.36
				(0.08)	(0.19)	(0.15)	(0.14)
$r_{\pi 0}$	Normal	1.8	0.1	1.78	1.81	1.72	1.66
				(0.08)	(0.10)	(0.10)	(0.09)
$r_{\pi 1}$	Normal	-0.3	0.1	-0.22	-0.22	-0.30	-0.39
				(0.09)	(0.12)	(0.10)	(0.09)
Implied parameters							
pseudo EIS: $\frac{1-h}{(1+h)\sigma_c}$				0.110	0.099	0.167	0.204
slope of PC: $\frac{(1-\beta\xi_p)(1-\xi_p)}{(1+\beta\gamma_p)\xi_p}$				0.011	0.007	0.012	0.018

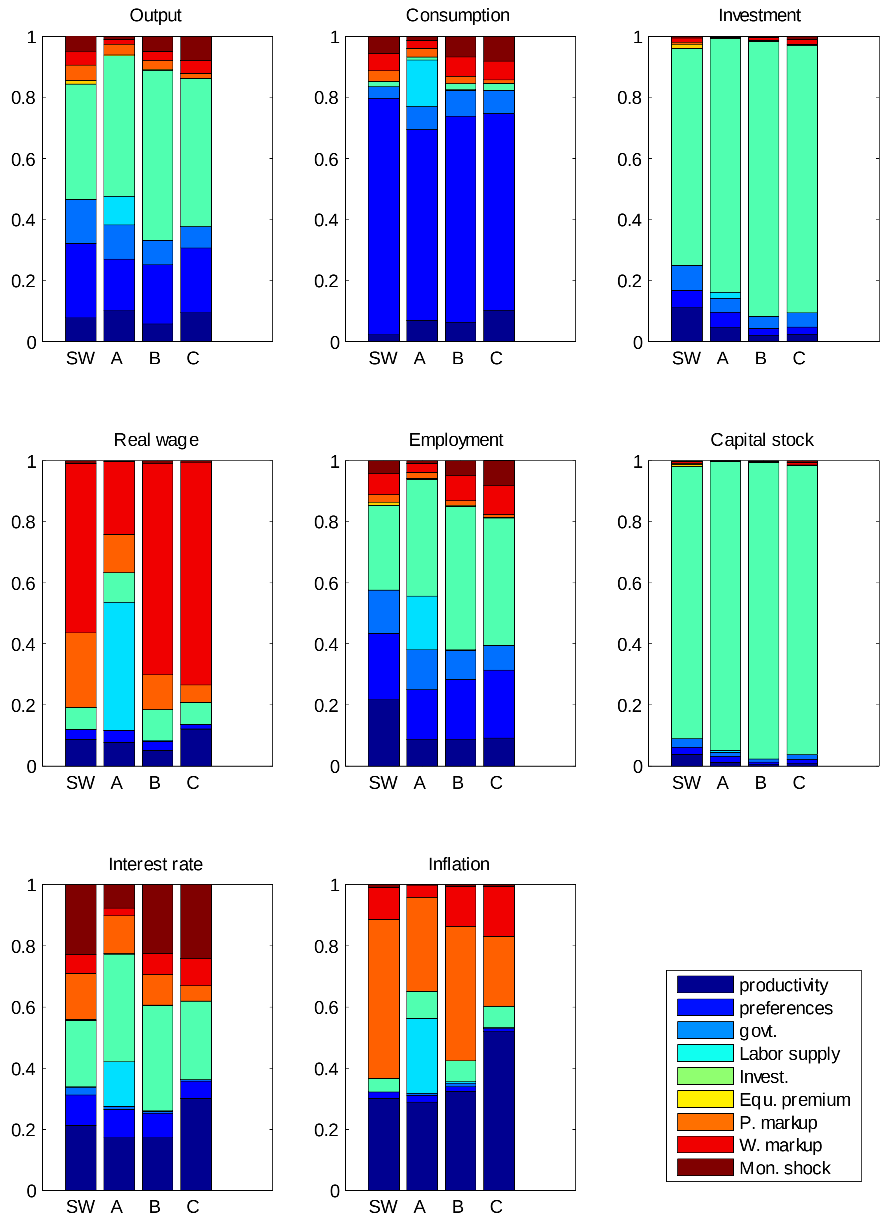
Estimated time series of capital and shocks



Estimated time series of shocks



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

- Building on Svensson (1999), Giannoni and Woodford (2002), Svensson and Woodford (2003, 2004)

- 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**

	var(π_t)	var(Y_t)	Loss
Case FI	2.0028	11.6477	3.6402
Case AI	2.6206	5.8146	4.4689

Avenues for future research

- Real time application with mixed frequencies