Short-term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility

Massimiliano Marcellino[1], Mario Porqueddu[2] and Fabrizio Venditti[2]

[1]EUI and Bocconi University

[2]Banca d'Italia - Research Department

Plan of the talk

- Motivation
- The model
- Estimation strategy
- An empirical application: forecasting euro area GDP
- Full sample results
- Daily business: some bayesian tools for nowcasting
- Out of sample: point and density forecast evaluation
- Some concluding remarks

Motivation 1

- Interest in policy making and forecasting in probability distributions around a central forecast
- Fan charts: Bank of England, Bank of Canada, Norges Bank, SA Reserve and Sveriges Riksbank, more recently Bank of Italy and also the US Fed (2008)
- Density forecasts are sensitive to shifts in the parameters of the model: Jore, Mitchell, and Vahey (2010), Clark (2011)
- Discrete breaks far away in the past: use sample split
- More recently: increasing interest in modeling small continuous breaks (time varying parameter models, Cogley and Sargent, 2003, Primiceri, 2005, large literature following)
- The Great Recession: spot light on volatility breaks (end to the Great Moderation?)

Motivation 2

- Most of the work on density forecast/time varying models falls in the medium/long term forecasting literature
- Nowcasting/Short term forecasting is a world of its own
  - mixed frequency data
  - ragged edge data
  - different timeliness (soft/hard data)
- There are existing tools that deal with the above issues but
  - No applications on density forecasts
  - Time constant parameters (some allow for discrete random breaks in the mean: MS models)
- Galvao (2009), STAR-MIDAS Guerin Marcellino (2011) MS-MIDAS Carriero, Clark, Marcellino (2012) U-MIDAS with stochastic volatility in the context of Nowcasting/Short term forecasting

Our contribution

- Extend the mixed frequency factor model by Mariano and Murasawa (2003) to account for continuous shifts in volatility
- Derive some interesting tools:
  - Density forecasts / fan charts for GDP short term forecasts
  - Probability distributions of the news content of indicator releases

Empirical application: forecasting euro area GDP

- We document a dramatic increase in both common and idiosyncratic business cycle volatility in the euro area in the past few years
- Evaluate the contribution to forecast accuracy of stochastic volatility in terms of:
  - Point forecast accuracy (RMSE)
    - S-vol lowers uniformly but marginally RMSE
  - Ability to produce normalized forecast errors (computed via PITS) which are close to normal
    - The model produces good pits with and without S-vol
  - Interval forecast accuracy (coverage rates)
    - S-vol improves significantly the coverage rates

The model

$$\begin{pmatrix} y^\star_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu^\star_1 \\ \mu_2 \end{pmatrix} + \beta f_t + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix}$$

$$y_{1t} = \frac{1}{3} y^\star_{1,t} + \frac{2}{3} y^\star_{1,t-1} + y^\star_{1,t-2} + \frac{2}{3} y^\star_{1,t-3} + \frac{1}{3} y^\star_{1,t-4}$$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \beta_1 (\frac{1}{3} f_t + \frac{2}{3} f_{t-1} + f_{t-2} + \frac{2}{3} f_{t-3} + \frac{1}{3} f_{t-4}) \\ \beta_2 f_t \end{pmatrix} + \begin{pmatrix} \frac{1}{3} u_{1,t} + \frac{2}{3} u_{1,t-1} + u_{1,t-2} + \frac{2}{3} u_{1,t-3} + \frac{1}{3} u_{1,t-4} \\ u_{2,t} \end{pmatrix}$$

Idiosyncratic shocks: the baseline model

$$
\begin{aligned}
f_t &= \sum_{j=1}^{p_f} \phi_j^f f_{t-j} + \epsilon_t^f & \epsilon_t^f \sim N(0, \sigma_f) \\
u_{1,t} &= \sum_{j=1}^{p_1} \phi_j^1 u_{1,t-j} + \epsilon_t^1 & \epsilon_t^1 \sim N(0, \sigma_1) \\
u_{2,t} &= \sum_{j=1}^{p_2} \phi_j^2 u_{2,t-j} + \epsilon_t^2 & \epsilon_t^2 \sim N(0, \sigma_2)
\end{aligned}
$$

Stochastic volatility

$$\begin{aligned} f_t &= \sum_{j=1}^{p_f} \phi_j^f f_{t-j} + \epsilon_t^f (\lambda_{f,t})^{0.5} & \epsilon_t^f \sim N(0,1) \\ u_{1,t} &= \sum_{j=1}^{p_1} \phi_j^1 u_{1,t-j} + \epsilon_t^1 (\lambda_{1,t})^{0.5} & \epsilon_t^1 \sim N(0,\sigma_1) \\ u_{2,t} &= \sum_{j=1}^{p_2} \phi_j^2 u_{2,t-j} + \epsilon_t^2 (\lambda_{2,t})^{0.5} & \epsilon_t^2 \sim N(0,\sigma_2) \end{aligned}$$

- We let the log-stochastic volatility components follow a random walk without drift

$$log(\lambda_{i,t}) = log(\lambda_{i,t-1}) + \eta_{i,t} \qquad \eta_{i,t} \sim N(0, \sigma_{\eta,i})$$

State space representation

- The model has a time varying state space representation

$$
\begin{aligned}
y_t &= F\mu_t \\
\mu_t &= H\mu_{t-1} + \eta_t \quad \eta_t \sim N(0, Q_t) \\
\Lambda_t &= \Lambda_{t-1} + \zeta_t \qquad \zeta_t \sim N(0, \Xi)
\end{aligned}
$$

- $\mathbf{y_t}$ collects both quarterly and monthly variables,
- μ_t includes the unobserved factor and idiosyncratic shocks
- Q_t collects the drifting volatilities $\sigma_i \lambda_{i,t}$

6 blocks of parameters: 6 steps Metropolis Hastings within Gibbs algorithm

$$\begin{aligned} y_t &= F\mu_t \\ \mu_t &= H\mu_{t-1} + \eta_t \quad \eta_t \sim N(0, Q_t) \\ \Lambda_t &= \Lambda_{t-1} + \zeta_t \quad \zeta_t \sim N(0, \Xi) \end{aligned}$$

1. Elements of F (β)
2. Elements of H (ϕ)
3. Time constant elements of Q_t (σ_i)
4. Time varying elements of Q_t ($\lambda_{i,t}$)
5. Variances of s-vol ($\sigma_{\eta,i}$)
6. The unobserved state vector (μ_t)

Uncorrelated disturbances: estimation can be performed equation by equation

Steps 1-2: β_i and σ_i

- Take a measurement equation

$$y_{i,t} = \beta_i f_t + u_{i,t}$$

- Autocorrelated $1 - \phi(L)$ and heteroscedastic $\lambda_{i,t}^{0.5}$ residuals
- Filter with $1 - \phi(L)$ and divide by $\lambda_{i,t}^{0.5}$
- f_t and all other parameters can be treated as known
- This is a standard regression
- Normal-gamma conjugate prior $\rightarrow$ Normal-gamma posterior

Step 3: ϕ_i

- Take a transition equation

$$\mu_{i,t} = \sum_{j=1}^{p_i} \phi_i \mu_{i,t-j} + \eta_{i,t}$$

- heteroscedastic $\lambda_{i,t}^{0.5}$ residuals
- divide by $\lambda_{i,t}^{0.5}$
- This is a standard regression
- Normal conjugate prior $\rightarrow$ Normal posterior
- Discard explosive roots

Step 4-5: $\lambda_{i,t}$, $\sigma_{\eta,i}$

- Use block-by-block Jacquier-Polson-Rossi algorithm
- Involves drawing from a candidate density (log-normal)
- Metropolis acceptance step

Step 6: μ_t

- Conditional on $F(\beta)$, $H(\phi)$, $Q_t(\sigma_i, \lambda_{i,t})$ use state space representation
- Durbin and Koopman disturbance smoother gives draws of μ_t
- Missing values in GDP equation are treated as in Mariano/Murasawa (skip the filtering step)

Priors

- Estimate with OLS on a training sample of τ initial observations
- Normal prior means are set at OLS estimates, variances at 10^3 the OLS variances
- Gamma degrees of freedom set to $\tau + 1$ for time constant variances
- Gamma degrees of freedom for the variance of $\lambda_{i,t}$ set to 1 and scale parameter to 0.025 (in line with Clark, 2011)
- We set τ to 36 (first three years of data).

Empirical application: Euro area GDP forecasting - indicators

Table: Variable selection summary

Indicator	Country
Quarterly series	
GDP	Euro Area
Monthly series	
Industrial Production	Euro Area
Industrial Production - Pulp/paper	Euro Area
Business Climate - IFO	Germany
Economic Sentiment Indicator	Euro Area
PMI composite	Euro Area
Exchange rate	US-Euro
10y spread	US-Euro
Michigan Consumer Sentiment	US

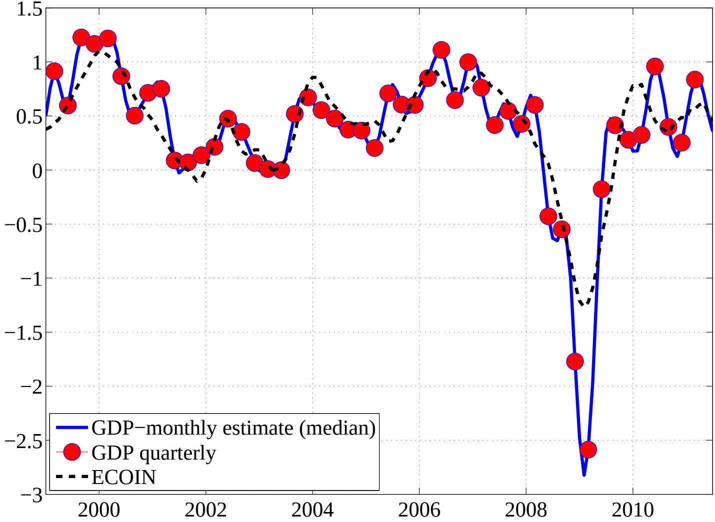
FULL SAMPLE RESULTS

Factor Loadings - posterior estimates

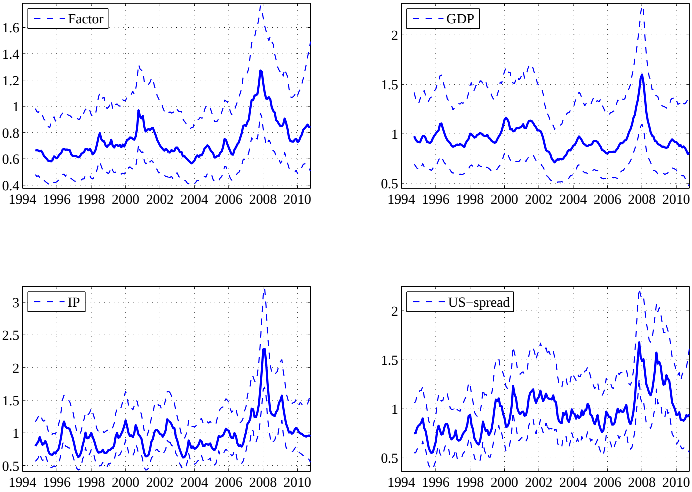
Table: Factor Loadings - posterior estimates

Percentiles	25th	50th	75th
GDP	0.27	0.38	0.54
IP	0.40	0.49	0.60
IP-PULP	0.23	0.29	0.36
IFO	0.10	0.12	0.13
ESI	0.10	0.12	0.14
PMI	0.12	0.13	0.15
US $ TO EURO	-0.08	-0.05	-0.02
US-spread	-0.06	-0.04	-0.02
Michigan Consumer	0.04	0.06	0.08

GDP: Median monthly estimate and Eurocoin



Stochastic volatilities - factor and selected indicators

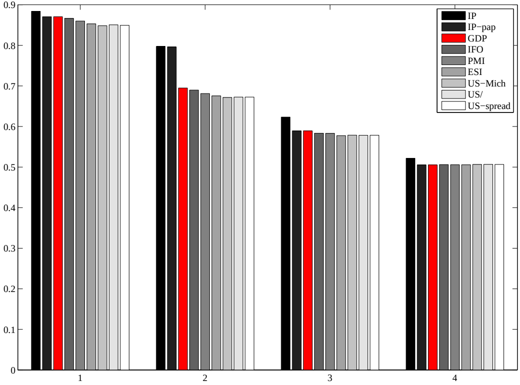


NEWS AND FORECASTS

Stylized data release calendar

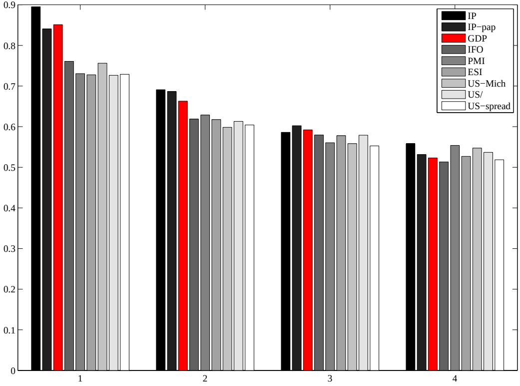
Indicator	Timing	Publication lag
IP	$11^{th} - 15^{th}$ of month	2
IP-PULP	$11^{th} - 15^{th}$ of month	2
GDP	1 day after IP	2
IFO	$20^{th} - 30^{th}$ of month	0
PMI	$20^{th} - 30^{th}$ of month	0
ESI	$20^{th} - 30^{th}$ of month	0
Michigan Consumer	Last Friday of the month	0
dollar-euro	Last day of month(Monthly ave.)	0
US-spread	Last day of month(Monthly ave.)	0

RMSE at different releases



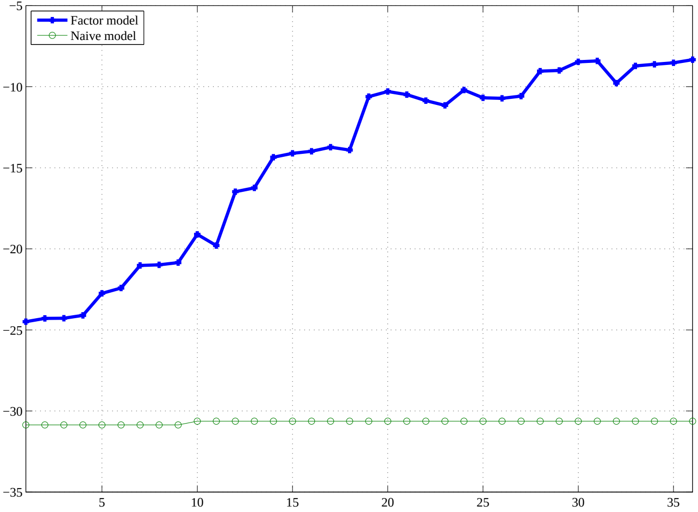
Note: ratio of the RMSE of the factor model with stochastic volatility to that of a naive constant growth model.

Forecast dispersion at different releases



Note: Standardized interquartile range (difference between the 75th and the 25th percentiles standardized by the median)

Log-predictive score at different releases

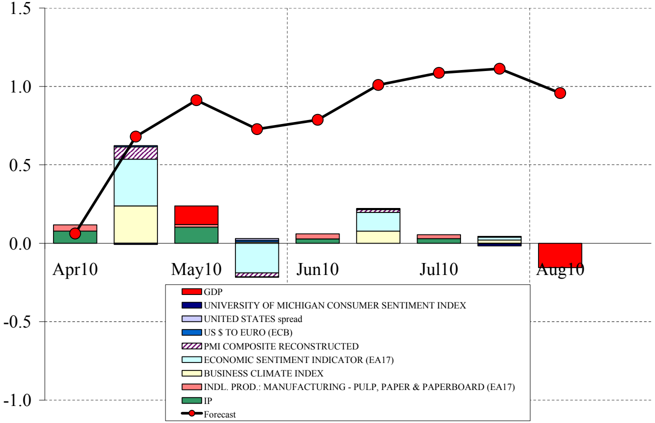


section

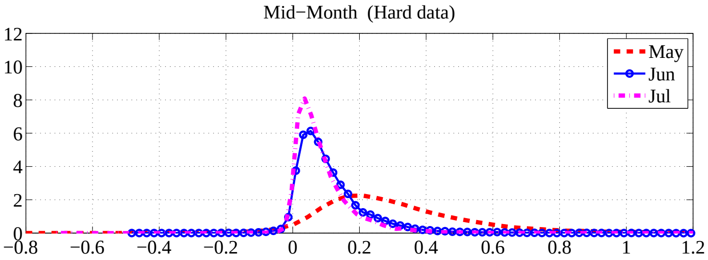
News content of data releases

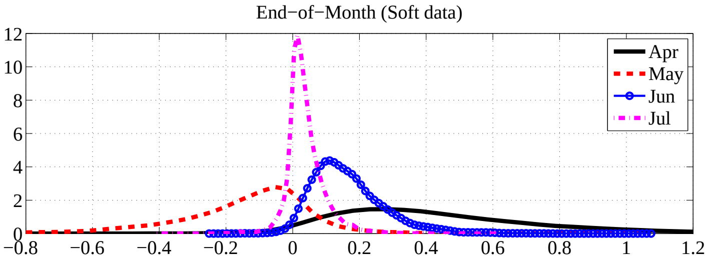
- Kalman smoother allows you to "dissect" the news content of each data release, taking into account the ragged-edge nature of monthly releases
- Various definitions of "news" in the literature: they all have flaws
- Recent contribution by Banbura-Modugno settles the issue. They show how to map monthly variables forecast errors into projection revisions
- We use their methodology to decompose the forecast revisions for 2010Q2 as new information accumulates

News and forecast evolution 2010Q2: median posterior estimate

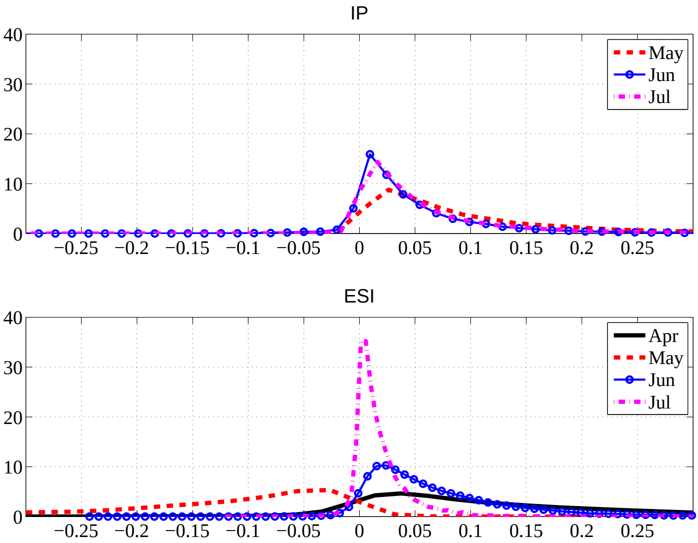


Our model assigns a probability to the overall revision...





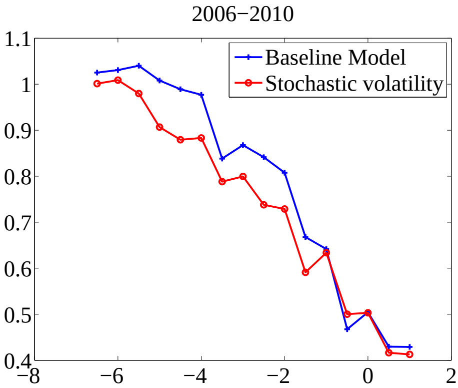
... and to the contributions!



section

FORECAST EVALUATION

Relative RMSE at different horizons



Coverage baseline

Table: Coverage Rates - Baseline Model

Nom Cov	Backcast		Nowcast		1 step ahead	
	Coverage	P-value	Coverage	P-value	Coverage	P-value
0.1	0.14	0.63	0.15	0.25	0.17	0.15
0.2	0.32	0.26	0.23	0.60	0.26	0.29
0.3	0.50	**0.08**	0.41	**0.08**	0.42	**0.05**
0.4	0.59	**0.09**	0.50	0.11	0.55	**0.02**
0.5	0.59	0.41	0.56	0.33	0.58	0.22
0.6	0.64	0.73	0.67	0.26	0.59	0.88
0.7	0.77	0.44	0.73	0.62	0.61	0.13
0.8	0.86	0.41	0.79	0.81	0.65	**0.01**
0.9	1.00	**0.09**	0.88	0.60	0.74	**0.01**

Coverage Stochastic Volatility

Table: Coverage Rates - Model with Stochastic Volatility

Nom Cov	Backcast		Nowcast		1 step ahead	
	Coverage	P-value	Coverage	P-value	Coverage	P-value
0.1	0.09	0.89	0.14	0.40	0.05	**0.04**
0.2	0.18	0.83	0.26	0.29	0.23	0.60
0.3	0.32	0.86	0.32	0.75	0.30	0.96
0.4	0.45	0.62	0.41	0.88	0.44	0.52
0.5	0.59	0.41	0.47	0.63	0.48	0.81
0.6	0.68	0.43	0.61	0.92	0.58	0.69
0.7	0.86	**0.04**	0.73	0.62	0.71	0.83
0.8	0.91	0.10	0.82	0.71	0.73	0.19
0.9	0.95	0.24	0.89	0.87	0.77	**0.02**

Some concluding remarks

- We introduce a mixed frequency factor model with stochastic volatility, and develop a Bayesian procedure for its estimation.
- We use it to model quarterly euro area GDP growth and a set of monthly indicators.
- In sample results show the relevance of changes in volatility. In addition, the estimated monthly GDP tracks very well the much more complex Eurocoin.
- We also show how, in a given quarter, the factor model can be used to assess the uncertainty around the news content of monthly releases of hard, soft and financial indicators.
- Finally, we evaluate out of sample point and density forecasts accuracy of the model, finding that SV improves substantially density forecasts

THANK YOU FOR THE ATTENTION