Forecasting German key macroeconomic variables using a large dataset

Inske Pirschel and Maik H. Wolters

Kiel University and Kiel Institute for the World Economy

14th IWH-CIREQ Macroeconometric Workshop 2013
Halle

3rd December 2013

C | A | U Christian-Albrechts-Universität zu Kiel
Wirtschafts- und Sozialwissenschaftliche Fakultät

ifw

Motivation

Forecasting Trade-off:

increase forecasting accuracy by including as much information as possible vs. dimensionality problems of large datasets

Solution:

aggregate information contained in large datasets to overcome dimensionality problem

Motivation

Forecasting Trade-off:

increase forecasting accuracy by including as much information as possible vs. dimensionality problems of large datasets

Solution:

aggregate information contained in large datasets to overcome dimensionality problem

Contribution

- study the performance of three alternative large scale approaches to forecast key macroeconomic variables
- apply methods to German economy

Contribution

- study the performance of three alternative large scale approaches to forecast key macroeconomic variables
- apply methods to German economy

Forecasting Models

3 large scale approaches:

- Large Bayesian Vectorautoregression (BVAR)

Forecasting Models

3 large scale approaches:

- Large Bayesian Vectorautoregression (BVAR)
- Equal Weighted/Bayesian Model Averaging (EWA, BMA)

Forecasting Models

3 large scale approaches:

- Large Bayesian Vectorautoregression (BVAR)
- Equal Weighted/Bayesian Model Averaging (EWA, BMA)
- Static/ Dynamic Factormodels (FAAR, FAVAR, DFM)

Forecasting Models

3 large scale approaches:

- Large Bayesian Vectorautoregression (BVAR)
- Equal Weighted/Bayesian Model Averaging (EWA, BMA)
- Static/ Dynamic Factormodels (FAAR, FAVAR, DFM)

Forecasting Models

5 small benchmark models:

- Iterative univariate Autoregression
- Direct univariate Autoregression
- Random Walk
- Vector Autoregression (with and without lag-selection criteria)
- Small Baysian Vector Autoregression

Forecasting Models

5 small benchmark models:

- Iterative univariate Autoregression
- Direct univariate Autoregression
- Random Walk
- Vector Autoregression (with and without lag-selection criteria)
- Small Baysian Vector Autoregression

A Large Dataset for Germany

- builds on Schumacher (2007)
- 124 variables in quarterly frequency

A Large Dataset for Germany

- builds on Schumacher (2007)
- 124 variables in quarterly frequency
- 5 data categories:
  1. Use of GDP and Gross value added
  2. Prices
  3. Manufacturing (turnover, production and received orders)
  4. Construction
  5. Surveys

A Large Dataset for Germany

- builds on Schumacher (2007)
- 124 variables in quarterly frequency
- 5 data categories:
  1. Use of GDP and Gross value added
  2. Prices
  3. Manufacturing (turnover, production and received orders)
  4. Construction
  5. Surveys
- sample period: 1978:1 - 2013:1

A Large Dataset for Germany

- builds on Schumacher (2007)
- 124 variables in quarterly frequency
- 5 data categories:
  1. Use of GDP and Gross value added
  2. Prices
  3. Manufacturing (turnover, production and received orders)
  4. Construction
  5. Surveys
- sample period: 1978:1 - 2013:1
- source: Datastream, Federal German Statistical Office
- seasonally adjusted (Census-X12), logs
- rescaling of West German data prior 1991

A Large Dataset for Germany

- builds on Schumacher (2007)
- 124 variables in quarterly frequency
- 5 data categories:
  1. Use of GDP and Gross value added
  2. Prices
  3. Manufacturing (turnover, production and received orders)
  4. Construction
  5. Surveys
- sample period: 1978:1 - 2013:1
- source: Datastream, Federal German Statistical Office
- seasonally adjusted (Census-X12), logs
- rescaling of West German data prior 1991

Setup

- compute forecasts for GDP growth, CPI inflation, 3 month money market rate, unemployment rate
- forecast horizon $h = 1, ..., 8$
- expanding window and rolling window forecasting scheme
- measure out-of-sample forecast accuracy in terms of relative root mean squared forecast errors (RMSFE)

Forecasting Models - Large BVAR

- VAR(p): $Y_t = c + A_1 Y_{t-1} + ... + A_p Y_{t-p} + \epsilon_t$
- prior on model parameters (Litterman 1996, Kadiyala and Karlson 1997):

$$E[(A_k)_{ij}] = \begin{cases} \delta_i = 1 \quad \text{or} \quad 0 & \text{for} \quad i = j, k = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$Var[(A_k)_{ij}] = \begin{cases} \frac{\lambda^2}{k^2} & \text{for} \quad i = j \\ \frac{\lambda^2 \sigma_i^2}{k^2 \sigma_j^2} & \text{otherwise} \end{cases} \tag{2}$$

- Bayesian shrinkage (Banbura et al. 2010): increase tightness of prior as number of variables in model increases
- set tightness of prior so that large BVAR has same in-sample fit as small BVAR

Forecasting Models - Large BVAR

- VAR(p): $Y_t = c + A_1 Y_{t-1} + ... + A_p Y_{t-p} + \epsilon_t$
- prior on model parameters (Litterman 1996, Kadiyala and Karlson 1997):

$$E[(A_k)_{ij}] = \begin{cases} \delta_i = 1 \quad \text{or} \quad 0 & \text{for} \quad i = j, k = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$Var[(A_k)_{ij}] = \begin{cases} \frac{\lambda^2}{k^2} & \text{for} \quad i = j \\ \frac{\lambda^2 \sigma_i^2}{k^2 \sigma_j^2} & \text{otherwise} \end{cases} \tag{2}$$

- Bayesian shrinkage (Banbura et al. 2010): increase tightness of prior as number of variables in model increases
- set tightness of prior so that large BVAR has same in-sample fit as small BVAR

Forecasting Models - EWA, BMA

- n simple models: $y_{t+h} = \rho_0 + \sum_{j=1}^{p} \rho_j y_{t-j} + \beta_i x_{it} + \epsilon_{it}$
- forecast of model i: $\hat{y}_{i,T+h}$
- final forecast: $\hat{y}_{T+h} = \omega^{-1} \sum_{i=1}^{n} \hat{y}_{i,T+h}$

Equal Weighted Averaging (Stock and Watson 2001)

- $\omega^{-1} = \frac{1}{n}$

Forecasting Models - EWA, BMA

- n simple models: $y_{t+h} = \rho_0 + \sum_{j=1}^{p} \rho_j y_{t-j} + \beta_i x_{it} + \epsilon_{it}$
- forecast of model i: $\hat{y}_{i,T+h}$
- final forecast: $\hat{y}_{T+h} = \omega^{-1} \sum_{i=1}^{n} \hat{y}_{i,T+h}$

Equal Weighted Averaging (Stock and Watson 2001)

- $\omega^{-1} = \frac{1}{n}$

Bayesian Model Averaging (Wright 2003)

- $\omega^{-1} = P(M_i)$ with $P(M_i)$ as posterior probability of model i
- assign prior that all models are equally likely to be true

Forecasting Models - EWA, BMA

- n simple models: $y_{t+h} = \rho_0 + \sum_{j=1}^{p} \rho_j y_{t-j} + \beta_i x_{it} + \epsilon_{it}$
- forecast of model i: $\hat{y}_{i,T+h}$
- final forecast: $\hat{y}_{T+h} = \omega^{-1} \sum_{i=1}^{n} \hat{y}_{i,T+h}$

Equal Weighted Averaging (Stock and Watson 2001)

- $\omega^{-1} = \frac{1}{n}$

Bayesian Model Averaging (Wright 2003)

- $\omega^{-1} = P(M_i)$ with $P(M_i)$ as posterior probability of model i
- assign prior that all models are equally likely to be true

Forecasting Models - FAAR, FAVAR, DFM

Static Factor Model (Stock and Watson, 2005)

- $y_{t+h} = \rho_0 + \sum_{j=1}^{p} \rho_j y_{t-j} + \sum_{i=1}^{r} \gamma_i z_{it} + \epsilon_t$
- $z_{it}{}_{i=1}^{r}$ as first r principal components of matrix $x_{it}{}_{i=1}^{n}$
- Bernanke et al. (2005): augment VAR with static factors

Dynamic factor model

- (Forni et al. 2000) extract factors by dynamic principal components analysis in the frequency domain

Forecasting Models - FAAR, FAVAR, DFM

Static Factor Model (Stock and Watson, 2005)

- $y_{t+h} = \rho_0 + \sum_{j=1}^{p} \rho_j y_{t-j} + \sum_{i=1}^{r} \gamma_i z_{it} + \epsilon_t$
- $z_{it}{}_{i=1}^{r}$ as first r principal components of matrix $x_{it}{}_{i=1}^{n}$
- Bernanke et al. (2005): augment VAR with static factors

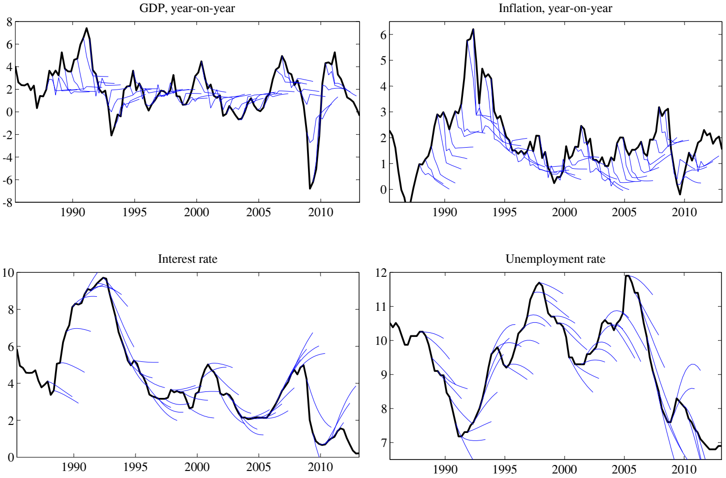
Dynamic factor model

- (Forni et al. 2000) extract factors by dynamic principal components analysis in the frequency domain

Results

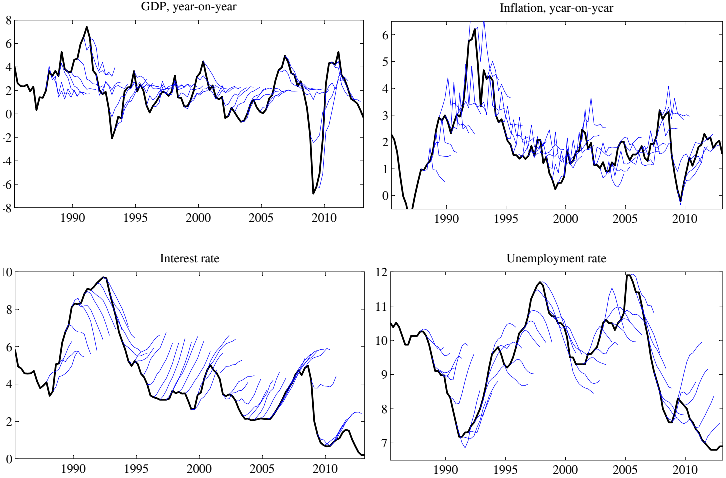
Results

Forecasting Performance - Large BVAR



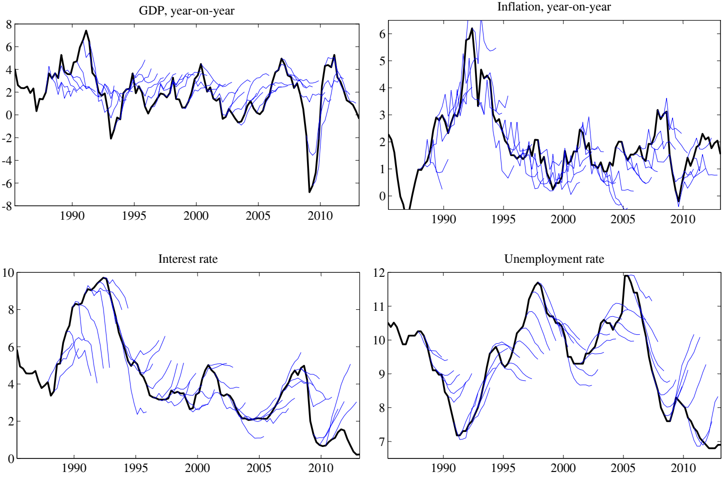
Results

Forecasting Performance - BMA



Results

Forecasting Performance - FAAR



Results

- quarter-on-quarter GDP growth time series shows very little persistence, returns to average growth rate after 1 - 2 quarters

$\rightarrow$ advantage for models which quickly return to steady state

- CPI inflation also not very persistent, however there are changes in the trend

$\rightarrow$ hard to capture for all models

Results

- quarter-on-quarter GDP growth time series shows very little persistence, returns to average growth rate after 1 - 2 quarters

$\rightarrow$ advantage for models which quickly return to steady state

- CPI inflation also not very persistent, however there are changes in the trend

$\rightarrow$ hard to capture for all models

- interest rate and unemployment show similar time series properties, both very persistent

$\rightarrow$ reflected in all model forecasts

Results

- quarter-on-quarter GDP growth time series shows very little persistence, returns to average growth rate after 1 - 2 quarters

$\rightarrow$ advantage for models which quickly return to steady state

- CPI inflation also not very persistent, however there are changes in the trend

$\rightarrow$ hard to capture for all models

- interest rate and unemployment show similar time series properties, both very persistent

$\rightarrow$ reflected in all model forecasts

Results

Forecasting Performance - RMFSE relative to random walk

GDP

horizon	LBVAR	BMA	FAAR
1	0.78**	0.85	0.83*
2	0.77**	0.79*	0.83*
3	0.77**	0.76**	0.79**
4	0.77**	0.75**	0.77*
5	0.72**	0.72*	0.74*
6	0.71**	0.69**	0.74**
7	0.71**	0.71**	0.75**
8	0.68**	0.67**	0.73**

Results

Forecasting Performance - RMFSE relative to random walk

CPI inflation

horizon	LBVAR	BMA	FAAR
1	0.83*	0.87**	0.87**
2	0.79**	0.82**	0.81**
3	0.84*	0.85**	0.85**
4	0.84	0.83**	0.85*
5	0.86**	0.81**	0.83**
6	0.88	0.83**	0.89*
7	0.87	0.86*	0.86**
8	0.84	0.82**	0.84**

Results

Forecasting Performance - RMFSE relative to random walk

Interest Rate

horizon	LBVAR	BMA	FAAR
1	0.81**	0.89*	0.91
2	0.82**	0.92	0.81*
3	0.85**	0.98	0.89
4	0.90	1.002	0.94
5	0.94	1.04	0.95
6	0.98	1.07	0.94
7	1.01	1.13	1.008
8	1.04	1.18	1.08

Results

Forecasting Performance - RMFSE relative to random walk

Unemployment Rate

horizon	LBVAR	BMA	FAAR
1	0.80**	0.79**	0.78**
2	0.79**	0.75**	0.79**
3	0.82*	0.78**	0.83
4	0.85	0.84	0.84
5	0.88	0.89	0.87
6	0.91	0.91	0.89
7	0.93	0.90	0.92
8	0.94	0.89	0.94

Results

- Large BVAR significantly dominates the random walk GDP forecast for all horizons at 5 % level
- for $h = 1, \ldots, 3$ Large BVAR also significantly outperforms random walk forecasts for CPI inflation, interest rate and unemployment at 5 or 10 % level

Results

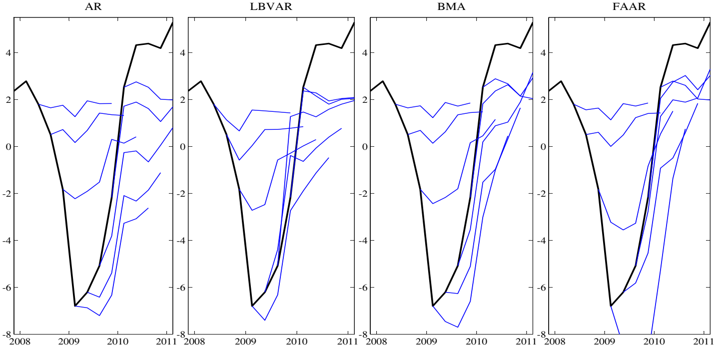
- Large BVAR significantly dominates the random walk GDP forecast for all horizons at 5 % level
- for $h = 1, ..., 3$ Large BVAR also significantly outperforms random walk forecasts for CPI inflation, interest rate and unemployment at 5 or 10 % level
- relative performance of BMA, FAAR, AR depends on the predicted variable

Results

- Large BVAR significantly dominates the random walk GDP forecast for all horizons at 5 % level
- for $h = 1, ..., 3$ Large BVAR also significantly outperforms random walk forecasts for CPI inflation, interest rate and unemployment at 5 or 10 % level
- relative performance of BMA, FAAR, AR depends on the predicted variable

Results

Forecasting Performance - Financial Crisis



Conclusion

- all models can provide a picture of current economic dynamics, but have difficulties predicting turning points and recessions
- Large BVAR shows a very good overall forecasting performance (especially for short horizons)

Conclusion

- all models can provide a picture of current economic dynamics, but have difficulties predicting turning points and recessions
- Large BVAR shows a very good overall forecasting performance (especially for short horizons)