NÁRODNÁ BANKA SLOVENSKA
EUROSYSTÉM

Short-term Forecasting of Real GDP Using Monthly Data

Juraj Huček, Alexander Karšay, Marián Vávra

Occasional Paper 1/2015

www.nbs.sk
Imricha Karvaša 1
813 25 Bratislva

research@nbs.sk

July 2015

ISSN 1337-5830

The views and results presented in this paper are those of the authors and do not necessarily represent the official opinions of the National Bank of Slovakia.

Short-term Forecasting of Real GDP Using Monthly Data[1]

Occasional paper NBS

Juraj Huček, Alexander Karšay, Marián Vávra[2]

Abstract

This occasional paper considers the problem of forecasting, nowcasting, and backcasting the Slovak real GDP growth rate using approximate factor models. Three different versions of approximate factor models are proposed. Forecast comparison with other models such as bridge equation models and ARMA models is also provided. Our results reveal that factor models clearly outperform an ARMA model and can compete with bridge models currently used at the Bank. Therefore, we tend to incorporate factor models into the regular forecasting process at the Bank. Finally, we hold the view that future research should be devoted to further improvements of bridge models since these models are simple to construct, easy to understand, and widely used in central banks.

JEL classification: C22, C38, C52, C53, E27
Key words: factor models, principal components, bridge equations, short-term forecasting, GDP.

[1] We would like to thank Peter Tóth and Miroslav Gavura for useful comments and interesting suggestions. Responsibility for any errors is ours alone.

[2] Marián Vávra, Research Department of the NBS, marian.vavra@nbs.sk; Juraj Huček, Economic and Monetary Analyses Department of the NBS, juraj.hucek@nbs.sk; Alexander Karšay, Economic and Monetary Analyses Department of the NBS, alexander.karsay@nbs.sk.

1. INTRODUCTION

Economic policy decisions are based on the expected development of key economic variables. However, some of these variables, for instance real GDP and other National Accounts indicators, are available to policy makers with a substantial delay (approximately 10 weeks). Therefore, early and accurate estimates of GDP figures is of the key importance for policy makers.

Since the influential work of Stock and Watson (2002a,b), factor modelling has received considerable attention in the forecasting literature. These models offer three main advantages over other econometric models and/or methods which make them popular. First, factor models can deal with many economic variables (predictors) without running into a dimensionality problem (see, e.g., Stock and Watson (2002a,b); Bai and Ng (2008)). However, one of the main problems with the current state of factor modelling is the selection of relevant variables (see Boivin and Ng (2006)). Second, this class of models can eliminate the effect of idiosyncratic shocks and measurement errors contaminating economic series. As a result, factors can yield better information about the state of the economy to policy makers (see, e.g., Chauvet (1998); Kapetanios (2004)). Third, empirical evidence suggests that factor models can (in some cases) outperform other forecasting models and/or methods (see Stock and Watson (2002b); Camba-Mendez and Kapetanios (2005); or Banerjee and Marcellino (2006)).

The main task of this occasional paper is twofold. First, we construct small approximate factor models for forecasting, nowcasting and backcasting the Slovak real GDP growth rate. Second, we compare the performance of approximate factor models with existing forecasting models routinely used at the NBS (i.e. bridge equation models and ARMA models).

The paper is organized as follows. A brief description of approximate factor models altogether with other competing forecasting models is given in Sections 2 and 3. A dataset, consisting of 70 monthly economic indicators, is described in Section 4. A quasi-real time evaluation of the forecast performance of the selected models is presented in Sections 5. Section 6 concludes and summarizes.

2. FACTOR MODELS

2.1 INTRODUCTION

Consider an $(N\times 1)$ vector of stationary monthly economic data collected in $\boldsymbol{x}_t = (X_{1,t}, \ldots, X_{N,t})$' for $t = 1, \ldots, T$.[3] Variables in $\boldsymbol{x}_t$ are assumed to be decomposed into two components: a com-

[3] Note that factor models can deal with mixed frequency data as well. However, since our main task is to construct a model for monthly updates of the real GDP figures, a mixed frequency approach is not considered here. The interested reader is referred to Armesto et al. (2010) for a discussion.

mon component (consisting of a small number of factors) and an idiosyncratic shock component

$$\boldsymbol{x}_t = \boldsymbol{\mu} + \boldsymbol{\Lambda} \boldsymbol{f}_t + \boldsymbol{u}_t, \tag{1}$$

where $\boldsymbol{\mu}$ is an $(N \times 1)$ vector of intercepts, $\boldsymbol{\Lambda}$ denotes an $(N \times k)$ loading matrix and $\boldsymbol{f}_t$ represents a $(k \times 1)$ vector of common factors such that $k << N$.[4] It is usually assumed that $\boldsymbol{u}_t \sim NID(\boldsymbol{0}, \boldsymbol{\Omega}_u)$, where $\boldsymbol{\Omega}_u$ is a (diagonal) variance-covariance matrix.[5] The law of motion of common factors is assumed to follow a (linear) VAR process given by

$$\boldsymbol{f}_t = \sum_{i=1}^{p} \boldsymbol{A}_i \boldsymbol{f}_{t-i} + \boldsymbol{\epsilon}_t, \tag{2}$$

where $\boldsymbol{A}_i$ are $(k \times k)$ parameter matrices and $\boldsymbol{\epsilon}_t$ is a $(k \times 1)$ error vector such that $\boldsymbol{\epsilon}_t \sim NID(\boldsymbol{0}, \boldsymbol{\Omega}_\epsilon)$, where $\boldsymbol{\Omega}_\epsilon$ is a variance-covariance matrix. In the last step, the estimated (monthly) factors are temporarily aggregated and linked to the quarterly real GDP growth rates using, for instance, a simple linear ARX model given by

$$y_\tau = c + \sum_{j=0}^{q} \boldsymbol{\beta}_j' \boldsymbol{f}^*_{\tau-j} + \phi y_{\tau-1} + \eta_\tau, \tag{3}$$

where y_τ denotes the real GDP growth rate at time τ, $\eta_t \sim NID(0, \sigma^2_\eta)$, and $\boldsymbol{f}^*_{\tau-j}$ represents the lagged aggregated common factors.[6]

Banbura et al. (2010, pp. 12-15) show that the system of equations (1)–(3) can be written into a linear state-space model. Doz et al. (2011), among others, show that this types of models can be estimated via a quasi maximum likelihood (QML) method.[7] The QML method of (large-scale) dynamic factor models proceeds in two-steps. In the first step, the latent common factors in (1) are approximated using principal components. In the second step, the whole state-space model is estimated using some numerical optimization procedure where partly observed or missing observations are approximated (updated) via a kalman filter.

In contrast to the above described dynamic factor model, we focus on the approximate factor model (AFM) here. There are three conceptual differences between DFM and AFM. First, the DFM works with common (hidden) factors, whereas the AFM is based on principal components calculated from the variance-covariance matrix of $\boldsymbol{x}_t$. Although principal components and common factors are not the same quantities, it can be shown that under mild regular conditions such that a number of observations T and a number of economic variables N tend to infinity and that idiosyncratic shocks are only (weakly) cross- and serially-correlated, Bai (2003) es-

[4]Note that $\boldsymbol{\Lambda}$ allows for a partition (structural restrictions) in order to distinguish among nominal, real, financial or global factors (see Banbura et al. (2010) among others).

[5]Note, however, that some weak cross- and serial-correlation of shocks is possible to incorporate into the model.

[6]The interested reader is referred to Drost and Nijman (1993) for details about temporal aggregation of stochastic processes.

[7]Note that some technical details of the above described state-space model are omitted for the sake of simplicity.

tablished consistency of principal components as an estimator of latent common factors.[8] As a result, dealing with factors and principal components is asymptotically equivalent. What is more, although assumptions about T and N may seem to be too strong for most of macroeconomic applications, Tanaka and Kurozumi (2012) found that a principal component estimator performs well even when N is small (which is exactly our case). Second, the ragged edges in data are fixed by a kalman filter in the case of DFM, whereas by ARMA models applied to individual variables in the case of AFM. Third, the DFM is estimated simultaneously using, for instance, the QML method, whereas all equations in the AFM are estimated separately (step by step). This fact significantly reduces the computational burden and improves the robustness of the approximate factor models.[9]

2.2 APPROXIMATE FACTOR MODELS IN A NUTSHELL

Approximate factor models are constructed in the following steps:

Step 0 Consider an $(N \times 1)$ vector of second-order stationary, short-range dependent, real-valued, balanced, and standardized economic indicators collected in $\boldsymbol{z}_t = (z_{1,t}, \dots, z_{N,t})'$, for $t \in \{1, \dots, T\}$.[10]

Step 1 Since economic variables are both cross- and serially-correlated, the principal components (approximate factors) are calculated from the (long-run) variance-covariance matrix of standardized economic variables (i.e. all variables have zero means and unit variances). Motivated by the literature on estimation of the long-run variance in the presence of weak dependence (see Newey and West (1987, 1994)), the following estimator is used

$$\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Gamma}}_0 + \sum_{j=1}^{h} w(j/h)(\hat{\mathbf{\Gamma}}_j + \hat{\mathbf{\Gamma}}_j^{'}), \tag{4}$$

where $w(\cdot)$ are the Bartlett weights, h is a real-valued bandwidth such that $h \to \infty$ and $h/T \to 0$ as $T \to \infty$, and $\hat{\mathbf{\Gamma}}_j = T^{-1} \sum_{t=j+1}^{T} \boldsymbol{z}_t \boldsymbol{z}_{t-j}^{'}$ denotes a sample vector autocovariance at lag j.

Step 2 Calculate the first k components $\hat{\boldsymbol{f}}_t = (\hat{f}_{1,t}, \dots, \hat{f}_{k,t})'$ from the long-run variance-covariance matrix $\hat{\mathbf{\Sigma}}$. Among the various methods available in the literature, the following rules for selecting k are popular in the literature: (i) the variance rule; (ii) the average-root rule; or (iii) the broken-stick rule. For a detailed discussion of these rules the reader is referred

[8] See also Stock and Watson (2002a) for a discussion.

[9] Our choice of focusing on the AFM is partly motivated by the results of Arnoštová et al. (2011) who found that approximate models (slightly) outperform many other forecasting models and/or methods used at the Czech National Bank for short-term forecasting of real GDP. Considering a high degree of similarities between the Czech and Slovak economies, we hold the view that approximate factor models might be a good starting point in factor modelling at the National Bank of Slovakia.

[10] A balanced panel of data is assumed here only for simplicity of exposition. See Section 4 for details how we deal with ragged edges in data.

to Jolliffe (2005, Ch. 6). Based on findings in Yamamoto (2015) who showed that structural instability in economic variables inflates a number of selected principal components (factors) and thus produces spurious (non-useful) factors. Therefore, only a moderate number of economic variables considered in our study (see Table 1) and set $k \in \{1, 2\}$.

Step 3 Since the principal components calculated from monthly data are uncorrelated each other (due to an orthogonality condition of the eigenvectors), we can use finite-order AR models for forecasting individual principal components $m \in \{1, \ldots, 9\}$ months ahead.[11] Note that m depends on the setup of the forecasting exercise. The lag order p is determined by means of the Bayesian information criterion (BIC), defined according to Method 1 of Ng and Perron (2005), with the maximum allowable order sets equal to $\bar{p} = \lfloor 8(T/100)^{1/4} \rfloor$.[12] The same method of balancing is used both for approximate factor and bridge equation models. Finally, a series of monthly principal components $\{\hat{\boldsymbol{f}}_t\}$ is temporarily aggregated to quarterly figures $\{\hat{\boldsymbol{f}}^*_\tau\}$.

Step 4 Finally, identify and estimate, a simple ARX model for the target variable (i.e. the real GDP growth rate)

$$y_\tau = c + \sum_{j=0}^{q} \boldsymbol{\gamma}'_j \hat{\boldsymbol{f}}^*_{\tau-j} + \phi y_{\tau-1} + +\beta d_\tau + \eta_\tau, \tag{5}$$

$y_{\tau-j}$ denotes the lagged dependent variable, $\hat{\boldsymbol{f}}^*_{\tau-j}$ represents a vector of calculated principal components at lag j, d_τ stands for a dummy variable taking 1 for $\tau = 2007Q4$, -1 for $\tau = 2008Q1$ and zero otherwise, η_τ is an error term. The lag order q are determined by means of the BIC according to Method 1 of Ng and Perron (2005), with the maximum allowable orders set equal to $\bar{q} = \lfloor 8(T/100)^{1/4} \rfloor$. The forecast of the real GDP figures is based on the estimated parameters from the above equation and predicted and aggregated principal components (see the previous step).

2.3 Selection of Variables

A practical question in factor modelling is how much data are really needed for forecasting the target variable (the real GDP growth rates in our case). Boivin and Ng (2006) found that extending the dataset by variables bearing little information about the target variable does not have to improve the forecast accuracy of factor models at all. For this reason, three different ways of determining an appropriate set of economic variables are considered here: (i) an expert judgement approach; (ii) a stepwise regression; and (iii) a least absolute shrinkage and selection approach.

[11] Note that the maximal forecast horizon of m depends on the specification of the approximate factor model, namely the lag order q in (2).

[12] $\lfloor A \rfloor$ denotes an integer part of A.

Expert Judgement: The indicators within this approach were selected in a way which aims to maximize pairwise correlations with GDP within six data categories of indicators (see Table 5). Indicators with strong cross-correlations within the same category were considered quasi-duplicates and only the ones with the strongest relation to GDP were finally selected, even if their weaker duplicates had a stronger correlation with GDP than the best performing indicators in other categories.

Stepwise Regression: Stepwise regression is a systematic method for adding and removing predictors from a regression model based on statistical significance (usually a standard F-test). The method starts with an initial model – a linear regression of y_t (i.e. the real GDP growth rate) on the vector of control variables (denoted as $\boldsymbol{w}_t$).[13] Then, in each iteration the p-value of the F-statistic is computed to test models with and without new predictors.[14] The method terminates once no single step improves the model fit.

Least Absolute Shrinkage and Selection Method Another method which has become popular for selecting the relevant economic variables is the least absolute shrinkage and selection operator (LASSO). The main advantage of this approach is that it performs the selection and shrinkage simultaneously. Let RSS be the sum of squared residuals from a regression of y_t (i.e. the real GDP growth rate) on regressors $x_{i,t}$, for $i = 1, \ldots, N$. The LASSO estimator is a solution to the following problem

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta}} RSS + \lambda \sum_{j=1}^{M} |\beta_j|,$$

where $\lambda > 0$ is a penalty term controlling for shrinkage, meaning a number of "useful" regressors and $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are parameter vectors. In our case, the λ parameter is set in such a way to get (approximately) 10 key economic variables for forecasting real GDP.

2.4 Empirical Results

Three versions of approximate factor models are considered in this paper, each of them differs according to how the key economic variables are selected. In the first factor model, denoted as "AFM 1", the selection of variables is purely on expert judgement. The final set of economic variables can be found in Table 1 (left-panel). In the third factor model, denoted as " AFM 3", the selection of key economic variables results from a union of selected variables from both the stepwise regression and the LASSO method. The final set of economic variables can be

[13]The control group consists of the following indicators: Industrial confidence indicator, Turnover in selected branches (sa, current prices), Industrial production (DE, sa), Industry confidence indicator (production expectations, sa), Ifo DE (expectations), Eurozone manufacturing PMI, composite index.

[14]The cut-off p-value is set to 0.10 in our case.

found in Table 1 (right-panel). Lastly, in the case of the second factor model, denoted as "AFM 2", the full set of 70 economic variables is split into two sub-sets according to their relationship to real GDP. In particular, we distinguish between coincident and leading indicators. As in the previous model, the final two sets of economic variables result from a union of the selected variables from both the stepwise regression and the LASSO method for each group. The final sets of economic variables can be found in Table 1 (middle-panel). The calculated principal components from all approximate factor models are presented in Figure 1.

Figure 1: Principal Components from Approximate Factor Models

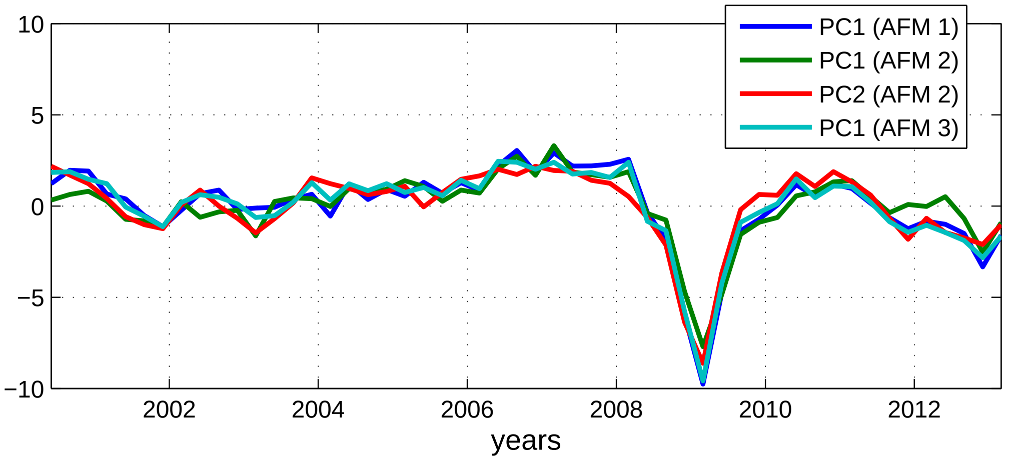


Table 1: List of Economic Variables for Selected Factor Models

AFM 1	AFM 2	AFM 3
Industrial confidence indicator	Turnover in selected branches, sa, current prices	Turnover in selected branches, sa, current prices
Consumer confidence indicator	Consumer confidence indicator	Industrial production, DE, sa
Spread 10y gov. bond yield vs 3m money market rate	Industrial confidence indicator	Major purchases over next 12 months
Number of registered unemployed, sa	Monthly employment, selected branches, sa	Cons. major purchases over next 12 m., EU
Industrial production index, sa	Industrial production, DE, sa	World Trade
Turnover in selected branches, sa, current prices	Consumer confidence indicator EU	Eurozone manuf. PMI; composite index
Export of goods, nominal, sa	Ifo DE	ESI EU
Industrial production, DE, sa		Industrial confidence indicator
CLI amplitude adusted EA		Goods trade balance, nominal, sa
Consumer confidence indicator EU	Eurozone manuf. PMI; composite index	Construction production, constant prices, sa
Real effective exchange rate, PPI	Manufacturing new orders, current prices, sa	Wages and Salaries
	Share Price Index, Germany	
	Cons. major purchases over next 12 m., EU	
	World Trade	
	Wages and Salaries	
	Ifo DE, expectations	
	Industry confidence indicator, production expectations, sa	

3. Other Forecasting Models

3.1 Bridge Equation Models

Bridge equation models represent a way of bridging monthly economic indicators with quarterly real GDP, whereas the monthly data is published earlier (see Angelini et al. (2011) for details). First of all, monthly data are forecasted until the end of the current quarter, typically using an ARMA model. Second, the monthly data is transformed to quarterly frequency and lastly enter the bridge equation with GDP, which is then estimated using the OLS method. The so called forecast combination has become very popular in the literature recently. In this approach, the estimated GDP growth rates from a larger number of various bridge equations are weighted using an accuracy criterion (e.g. AIC or RMSE, etc.). In our study, we refer to this GDP growth estimation method simply as the "bridge equation approach" and we use it as a benchmark to compare with the factor model results.

Formally, a bridge equation can be written as follows

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 x_{t-L} + \beta_3 d_t + \epsilon_t, \tag{6}$$

y_t is the real GDP growth rate at time t, $\beta_0, \beta_1, \beta_2, \beta_3$ are the unknown parameters estimated by the OLS method, x_{t-L} is the appropriately lagged stationary explanatory variable (e.g. the growth rate of industrial production or the unemployment rate, etc), d_t stands for a dummy variable taking 1 for $t = 2007Q4$, -1 for $t = 2008Q1$ and zero otherwise, and ϵ_t is an error term.

With monthly data available and GDP still missing for a given quarter, we can make use of such equation to obtain the GDP growth forecast (or nowcast). We will thus have a number of estimated GDP growth rates available, given that we are using a set of bridge equations. The method of forecast weighting requires the use of appropriately chosen weights. For this purpose, we apply an indicator of explanatory power of each bridge equation based on the AIC.[15] The lower the AIC value the better the fit of the bridge equation. The weighting scheme therefore has to reflect the need for a negative relation between AIC and the weight assigned to the given bridge equation forecast. We opted for a weighting scheme according to the study by Drechsel and Scheufele (2010) given by

$$w_i = \frac{\exp(-0.5(AIC_i - AIC_{min}))}{\sum_{i=1}^{N} \exp(-0.5(AIC_i - AIC_{min}))},$$

where AIC_i represents the fit of the ith bridge equation, AIC_{min} is the minimum of AIC values over N bridge models (indicators) applied in the analysis. Therefore, the better the fit the higher the weight assigned to the GDP growth forecast derived from the monthly indicator i.

[15] Alternative criteria such as RMSE or R^2 have been tested as well but no significant differences have been observed.

We consider three versions of bridge models in our study. First, it is a 40 variable (equation) model, denoted as "BEM 2". It includes monthly indicators of the Slovak economy, such as business surveys, industrial production and orders, interest rates and spreads and labour market variables (see Table 2). This type of model is also currently assisting the short-term forecasts at the NBS. Secondly, it is an 11-variable bridge model, denoted as "BEM 1", with indicators selected on expert judgement (as in the case of AFM 1, see the left-panel of Table 1). Finally, there is also a statistic bridge model, denoted as "BEM 3)", for which its 15 indicators were selected on the basis of more complex statistical methods, such as stepwise regressions and so-called LASSO method (as in the case of AFM 2, see the middle-panel of Table 1).

Table 2: List of Economic Variables for Bridge Equation Model: BEM 2

Industry confidence indicator, order books, sa	Major purchases over next 12 months
Industry confidence indicator, export order books, sa	Savings at present
Industry confidence indicator, stocks of finished products, sa	Savings over next 12 months
Industry confidence indicator, production expectations, sa	Statement on financial situation of household
Industry confidence indicator, selling price expectations, sa	ECB/NBS base rate
Industry confidence indicator, Employment expectations , sa	Euribor/bribor
Industrial confidence indicator	10y gov. bond yield (convergence criterion)
Consumer confidence indicator	Spread 10y gov. bond yield vs base rate
Retail confidence indicator	Spread 10y gov. bond yield vs 3m money market rate
Construction confidence indicator	CLI amplitude adusted
Services confidence indicator	Manufacturing new orders, current prices, sa
Economic sentiment indicator	Capital goods, new orders, current prices, sa
Consumers financial situation past 12m, sa	Consumer goods, new orders, current prices, sa
Consumers financial situation over next 12 months, sa	Industrial production, intermediate goods, sa
Consumers general economic situation over last 12 months, sa	Number of registered unemployed, sa
Consumers general economic situation over next 12 months	Industrial production index, sa
Price trends over last 12 months	Monthly employment, selected branches, sa
Price trends over next 12 months	HICP level
Unemployment expectations over next 12 months	HICP excluding energy level
Major purchases at present	Average monthly wages in selected branches, sa

3.2 ARMA Models

We also consider a finite-order ARMA(P,Q) model as a benchmark model for comparison with factor models and bridge equation models. The lag orders P and Q are determined by means of the BIC according to Method 1 of Ng and Perron (2005), with the maximum allowable orders set equal to $\bar{P} = \bar{Q} = \lfloor 8(T/100)^{1/4} \rfloor$.[16] The selected optimal lag orders are $P = 1$ and $Q = 0$.[17] Since the automatic lag order selection procedure does not guarantee the desirable properties of the estimated residuals for forecasting (e.g. no-serial correlation), the residuals from all estimated ARMA models have been inspected by standard diagnostic tests for serial correlation, heteroscedasticity and normality. No significant difference among a set of the estimated ARMA models have been observed.[18]

4. Data

We employ a set of 70 monthly economic indicators spanning the period January 2000 – December 2013. The dataset consists of the following six data categories: (A) Financial (12 series); (B) Output and activity (10 series); (C) Labour and Wages (4 series); (D) Prices (4 series); (E) Trade (4 series); (F) Domestic Surveys (23 series); and (G) Foreign Surveys (13 series). All relevant indicators are seasonally adjusted and transformed to assure stationarity (see Table 5 for details). Following Barhoumi et al. (2008), each variable is transformed according to one of the following three rules: (i) the three-months differences $X_{i,t} - X_{i,t-3}$ (denoted as "Δ"); (ii) the three-months log-differences $\log X_{i,t} - \log X_{i,t-3}$ (denoted as "$\Delta \log$"); and (iii) no transformation (denoted as "$-$"). All transformed series are normalized to have zero means and unit variances.

It is worth remarking that although the month-on-month changes might be preferred from a statistical point of view, the three-month changes have two advantages. First, the noise-to-signal ratio is reduced which implies that the extracted principal components have a higher explanatory power and reduces a number of components required in the analysis. Second, this type of transformation is convenient for a temporal aggregation of monthly data.

One of the major operational problems in multiple time series analysis is unbalanced datasets due to non-synchronous flow of data. In contrast to (pure) dynamic factor models where the missing observations in data are replaced by the "optimal" predictions of the factor(s) updated iteratively via a kalman filter (see Cuevas and Quilis (2012))[19], we deal with ragged edges by

[16] $\lfloor A \rfloor$ denotes an integer part of A.

[17] The results are available upon request from the authors.

[18] It is well known statistical fact that standard estimators of ARMA models suffer from a small sample bias (see, e.g., Yamamoto and Kunitomo (1984) and Engsted and Pedersen (2014)). However, Kim and Durmaz (2012) show that a bootstrap bias correction does not necessarily improve the forecast performance. The reason is that although the bootstrap procedure reduces the bias, it tends to increase the variance, and thus, the impact on the MSFE is ambiguous. Therefore, the bootstrap bias correction of the estimated AR parameters is not considered here.

[19] An advantage of this approach is that the estimated factors incorporate both cross-sectional and time-series

filling the missing observations of monthly indicators by predictions from a finite-order AR(p) model.[20] The same method of balancing is used both for approximate factor and bridge equation models.

5. Forecast Evaluation

The forecast performance of 7 models (3 approximate factor models, 3 bridge equation models, and an AR model) is evaluated using "quasi" real-time setup as in Barhoumi et al. (2008). This approach uses only the last available vintage of real GDP and monthly indicators (January 26, 2014) and imitates a time lag of monthly indicators according to Table 5. For the purpose of the forecast evaluation, we make an operational assumption that all models are run in the third week of each month to get updates of real GDP forecasts, nowcasts, and backcasts. The model forecasts are then evaluated in terms of the mean squared forecast error, forecast bias, and variability of forecasts. The results are presented in Tables 3 – 4. Individual forecasts of the real GDP growth rates over 9 consecutive months (i.e. forecasting, nowcasting, and backcasting) for a period 2010Q1–2013Q4 are depicted in Figure 2. Note that the forecast window consists of only 16 observations which does not allow for reliable inference using, for instance, the Diebold-Mariano test (see Diebold and Mariano (1995) and Vávra (2015)).[21]

Table 3: Forecast Evaluation: MSFE

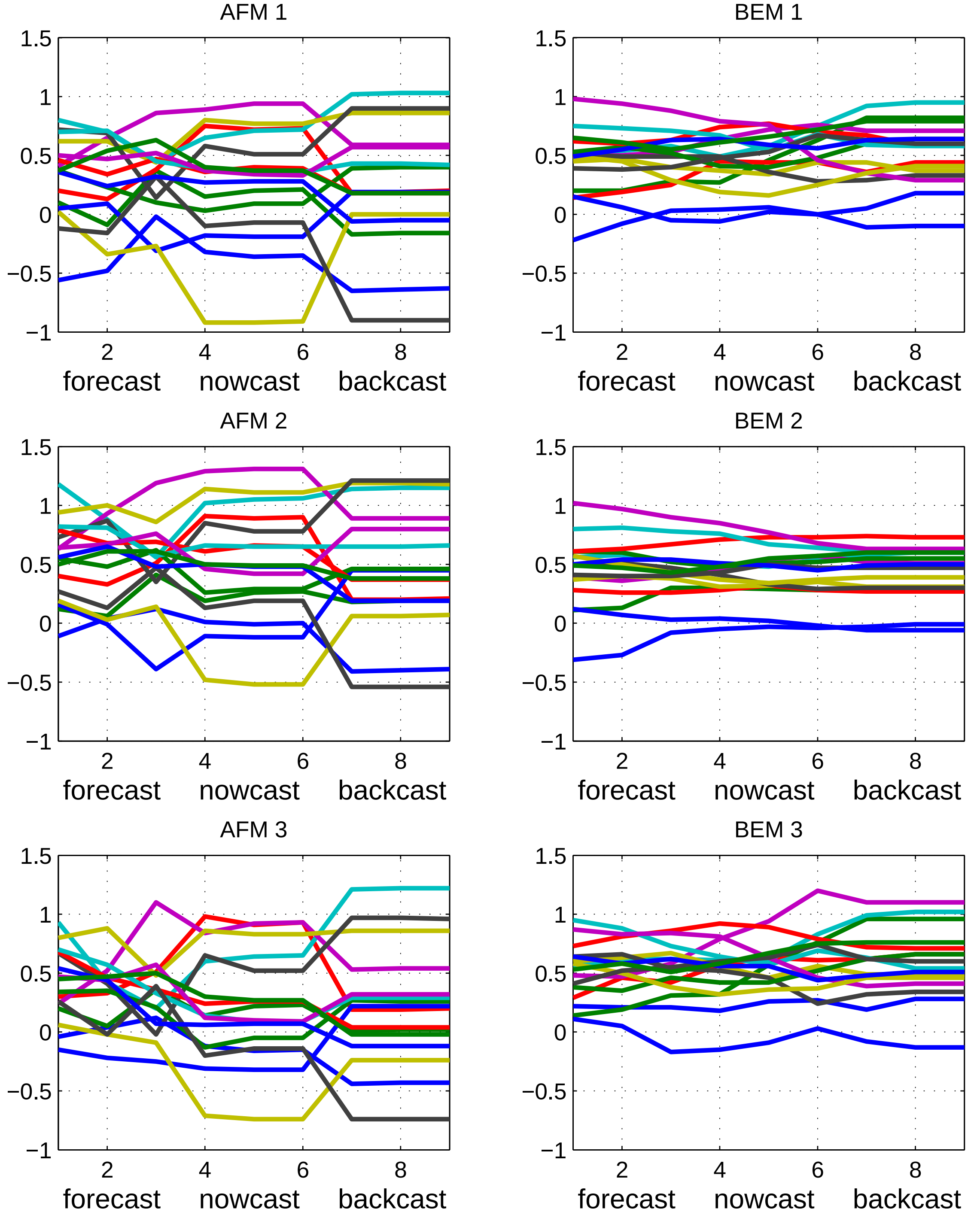
	horizon	ARMA	AFM 1	AFM 2	AFM 3	BEM 1	BEM 2	BEM 3
	1	0.47	0.22	0.38	0.25	0.27	0.28	0.31
forecasting	2	0.47	0.21	0.38	0.17	0.26	0.27	0.32
	3	0.47	0.18	0.36	0.18	0.25	0.25	0.32
	4	0.46	0.28	0.47	0.25	0.26	0.23	0.33
nowcasting	5	0.46	0.28	0.47	0.25	0.26	0.23	0.34
	6	0.46	0.28	0.47	0.25	0.30	0.22	0.41
	7	0.44	0.31	0.46	0.29	0.33	0.22	0.42
backcasting	8	0.44	0.31	0.46	0.29	0.33	0.22	0.43
	9	0.44	0.31	0.46	0.29	0.33	0.22	0.43
average		0.46	0.26	0.43	0.25	0.29	0.24	0.37

properties in the dataset. However, it does not work properly for models with just one factor (or a small number of factors) since is covers usually less than 50 % of variability of the indicators. Therefore, we are of the opinion that improvements in balancing can be achieved via simple univariate ARMA models applied to individual indicators with missing values.

[20] The lag order p is determined by means of the Bayesian information criterion (BIC), defined according to Method 1 of Ng and Perron (2005), with the maximum allowable order sets equal to $\bar{p} = \lfloor 8(T/100)^{1/4} \rfloor$, where $\lfloor A \rfloor$ denotes an integer part of A.

[21] Note that a limited window used for a forecast evaluation (i.e. 16 quarters) is determined mainly by a limited number of observations available. It is worth noting that there is no clear-cut on how many observations to reserve for a forecast evaluation in the literature (see Pesaran and Timmermann (2007) for a discussion). One should keep in mind a trade-off between a number of in-sample observations required for the estimation of model parameters and a number of out-of-sample observations necessary for a reasonable forecast evaluation. In our case, a forecast evaluation is based on approximately 30 % of available observations, the ratio which is in line with many other forecasting-based studies (see, e.g., Stock and Watson (1996), Liu and Jansen (2007), or Bai and Ng (2008)).

Figure 2: Forecast Errors of the Real GDP Growth Rates from Factor and Bridge Models



Note: Individual lines in figures represent forecast errors of the real GDP growth rates from the factor and bridge models over 9 consecutive months (i.e. forecasting, nowcasting, and backcasting) for a given period 2010Q1–2013Q4.

Table 4: Forecast Evaluation: Bias

	horizon	ARMA	AFM 1	AFM 2	AFM 3	BEM 1	BEM 2	BEM 3
forecasting	1	0.63	0.31	0.52	0.40	0.45	0.44	0.51
	2	0.63	0.27	0.51	0.32	0.45	0.44	0.52
	3	0.63	0.30	0.50	0.30	0.45	0.44	0.51
nowcasting	4	0.62	0.26	0.50	0.22	0.45	0.43	0.52
	5	0.62	0.26	0.50	0.21	0.46	0.43	0.53
	6	0.62	0.26	0.50	0.22	0.49	0.42	0.58
backcasting	7	0.61	0.23	0.45	0.21	0.50	0.42	0.58
	8	0.61	0.24	0.45	0.21	0.52	0.42	0.58
	9	0.61	0.24	0.45	0.21	0.52	0.42	0.58
average		0.62	0.26	0.49	0.26	0.48	0.43	0.55

The results suggest the following:

(i) AFM 3 and BEM 2 perform best in terms of the MSFE measure. Both models produce almost identical results and visibly lower MSFEs as compared to the estimated ARMA model.

(ii) By comparing the forecast performance of the factors models (i.e. AFM 1 - AFM 3), it can be concluded that the selection of economic variables matters.

(iii) In terms of a bias, some factors models (namely AFM 1 and AFM 3) produce forecasts with a significantly smaller bias as compared to the ARMA benchmark and even the bridge models over all forecast horizons. This favourable feature is, unfortunately, completely offset by high variability of factor model forecasts (see Figure 2 for details).

All in all, it can be concluded that since it is relatively easy to reduce a bias[22] of forecasts, our results slightly favour the bridge equation approach as compared to factor models (i.e. BEM 2 model).

Initial practical experience with a forecasting application of the aforementioned models suggests that they provide estimates of quarterly growth rates which are biased relative to actual GDP growth rates. This fact is broadly in line with formal evaluation results in Table 4. At present, a way of obtaining more precise estimates closer to actual GDP figures seems to be the application of the innovation (i.e. difference) in the estimated GDP growth rate for the current quarter with respect to the previous quarter estimate. This innovation is then added to the last available actual GDP growth rate, which results in a less biased and more precise GDP estimate (forecast, nowcast or backcast). This approach, however, will be subject to further testing in real time.

[22] Note that a forecast bias, if it is stable over time and/or over forecast horizons, can be easily reduced by just a constant from factor model forecasts.

6. Conclusion

In this paper we have evaluated the forecast performance of 3 different types of models often used for short-term forecasting of the real GDP growth rates. Overall, both bridge and factor models do produce visibly better results as compared to the ARMA benchmark. However, when comparing factor and bridge models, no single model clearly outperforms the others at all horizons and forecast measures. Nevertheless, it seems to be the case that bridge models (due to a possible bias reduction) may offer an interesting advantage over factor models.

References

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011). Short-term forecasts of euro area gdp growth. *The Econometrics Journal 14*, 25–44.

Armesto, M., K. Engemann, and M. Owyang (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review 92*, 521–36.

Arnoštová, K., D. Havrlant, L. Růžička, and P. Tóth (2011). Short-term forecasting of czech quarterly gdp using monthly indicators. *Czech Journal of Economics and Finance 6*, 566–583.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica 71*, 135–171.

Bai, J. and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146*, 304–317.

Banbura, M., D. Giannone, and L. Reichlin (2010). Nowcasting. *ECB Working Paper Series 1275*.

Banerjee, A. and M. Marcellino (2006). Are there any reliable leading indicators for us inflation and gdp growth? *International Journal of Forecasting 22*, 137–151.

Barhoumi, K., S. Benk, R. Cristadoro, A. Reijer, A. Jakaitiene, P. Jelonek, A. Rua, G. Rünstler, K. Ruth, and C. Nieuwenhuyze (2008). Short-term forecasting of gdp using large monthly datasets–a pseudo real-time forecast evaluation exercise. *ECB Working Paper 84*.

Boivin, J. and S. Ng (2006). Are more data always better for factor analysis? *Journal of Econometrics 132*, 169–194.

Camba-Mendez, G. and G. Kapetanios (2005). Forecasting euro area inflation using dynamic factor measures of underlying inflation. *Journal of Forecasting 24*, 491–503.

Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switching. *International Economic Review 39*, 969–996.

Cuevas, A. and E. Quilis (2012). A factor analysis for the spanish economy. *SERIEs 3*, 311–338.

Diebold, F. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business & economic statistics 13*, 253–263.

Doz, C., D. Giannone, and L. Reichlin (2011). A two-step estimator for large approximate dynamic factor models based on kalman filtering. *Journal of Econometrics 164*, 188–205.

Drechsel, K. and R. Scheufele (2010). Should we trust in leading indicators? evidence from the recent recession. Technical report.

Drost, F. and T. Nijman (1993). Temporal aggregation of garch processes. *Econometrica: Journal of the Econometric Society 61*, 909–927.

Engsted, T. and T. Pedersen (2014). Bias-correction in vector autoregressive models: A simulation study. *Econometrics 2*, 45–71.

Jolliffe, I. (2005). *Principal Component Analysis*. Wiley.

Kapetanios, G. (2004). A note on modelling core inflation for the uk using a new dynamic factor estimation method and a large disaggregated price index dataset. *Economics Letters 85*, 63–69.

Kim, H. and N. Durmaz (2012). Bias correction and out-of-sample forecast accuracy. *International Journal of Forecasting 28*, 575–586.

Liu, D. and D. Jansen (2007). Macroeconomic forecasting using structural factor analysis. *International Journal of Forecasting 23*, 655–677.

Newey, W. and K. West (1987). West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Newey, W. and K. West (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies 61*, 631–653.

Ng, S. and P. Perron (2005). A note on the selection of time series models. *Oxford Bulletin of Economics and Statistics 67*, 115–134.

Pesaran, M. and A. Timmermann (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics 137*, 134–161.

Stock, J. and M. Watson (1996). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. *National Bureau of Economic Research Working Paper 6607*.

Stock, J. and M. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167–1179.

Stock, J. and M. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*, 147–162.

Tanaka, S. and E. Kurozumi (2012). Investigating finite sample properties of estimators for approximate factor models when n is small. *Economics Letters 116*, 465–468.

Vávra, M. (2015). On a bootstrap-based test for forecast evaluations. *National Bank of Slovakia Working Paper 5*.

Yamamoto, T. and N. Kunitomo (1984). Asymptotic bias of the least squares estimator for multivariate autoregressive models. *Annals of the Institute of Statistical Mathematics 36*, 419–430.

Yamamoto, Y. (2015). Forecasting with non-spurious factors in us macroeconomic time series. *Journal of Business & Economic Statistics* (forthcoming).

A. TABLES

Table 5: Data Description

Type	Series	Delay (weeks)	Transformation	Source
Survey	Industry confidence indicator, order books, sa	0	–	EC
Survey	Industry confidence indicator, export order books, sa	0	–	EC
Survey	Industry confidence indicator, stocks of finished products, sa	0	–	EC
Survey	Industry confidence indicator, production expectations, sa	0	–	EC
Survey	Selling price expectations, sa	0	–	EC
Survey	Employment expectations for the months ahead, sa	0	–	EC
Survey	Industrial confidence indicator	0	–	EC
Survey	Consumer confidence indicator	0	–	EC
Survey	Retail confidence indicator	0	–	EC
Survey	Construction confidence indicator	0	–	EC
Survey	Economic sentiment indicator	0	–	EC
Survey	Consumers financial situation past 12m, sa	0	–	EC
Survey	Consumers financial situation over next 12 months, sa	0	–	EC
Survey	Consumers general economic situation over last 12 months, sa	0	–	EC
Survey	Consumers general economic situation over next 12 months	0	–	EC
Survey	Price trends over last 12 months	0	–	EC
Survey	Price trends over next 12 months	0	–	EC
Survey	Unemployment expectations over next 12 months	0	–	EC
Survey	Major purchases at present	0	–	EC
Survey	Major purchases over next 12 months	0	–	EC
Survey	Savings at present	0	–	EC
Survey	Savings over next 12 months	0	–	EC
Survey	Statement on financial situation of household	0	–	EC
Financial	ECB base rate	0	Δ	ECB
Financial	Euribor/Bribor	0	Δ	EMMI
Financial	10y gov. bond yield (convergence criterion)	2	Δ	Eurostat
Financial	Spread 10y gov. bond yield vs base rate	2	–	See above
Financial	Spread 10y gov. bond yield vs 3m money market rate	2	–	See above
Financial	Spread euribor vs base rate	0	–	See above
Output	CLI amplitude adusted	7	–	OECD

Output	Manufacturing new orders, current prices, sa	7	$\Delta \log$	SO SR
Output	Capital goods, new orders, current prices, sa	7	$\Delta \log$	SO SR
Output	Consumer goods, new orders, current prices, sa	7	$\Delta \log$	SO SR
Output	Industrial production, intermediate goods, sa	6	$\Delta \log$	SO SR
Labour market	Number of registered unemployed, sa	4	$\Delta \log$	Labour office
Output	Industrial production index, sa	6	$\Delta \log$	SO SR
Labour market	Monthly employment, selected branches, sa	7	$\Delta \log$	SO SR
Prices	HICP level	3	$\Delta \log$	SO SR
Prices	HICP excluding energy level	3	$\Delta \log$	SO SR
Labour market	Average monthly wages in selected branches, sa	7	$\Delta \log$	SO SR
Output	Turnover in selected branches, sa, current prices	7	$\Delta \log$	SO SR
Output	Construction production, constant prices, sa	6	$\Delta \log$	SO SR
Trade balance	Import of goods, nominal, sa	6	$\Delta \log$	NBS
Trade balance	Export of goods, nominal, sa	6	$\Delta \log$	NBS
Trade balance	Goods trade balance, nominal, sa	6	Δ	NBS
Prices	Producer price index, domestic market, nsa	5	$\Delta \log$	Eurostat
Survey	ESI EU	0	–	EC
Survey	Ifo DE	0	–	CESIfo
Output	Industrial production, DE, sa	7	$\Delta \log$	Eurostat
Output	CLI amplitude adjusted, EA	7	–	OECD
Survey	Consumer confidence indicator EU	0	–	EC
Financial	Share Price Index, Germany	3	$\Delta \log$	OECD
Commodities	Oil price, Brent, nsa, EUR	0	$\Delta \log$	ECB SDW
Financial	USD/EUR	0	$\Delta \log$	ECB SDW
Financial	Nominal effective exchange rate	1	$\Delta \log$	NBS
Financial	Real effective exchange rate, CPI	5	$\Delta \log$	NBS
Financial	Real effective exchange rate, PPI	5	$\Delta \log$	NBS
Financial	Real effective exchange rate, PPI manuf.	5	$\Delta \log$	NBS
Survey	Ifo DE, expectations	0	–	CESIfo
Survey	Order books, industry, EU	0	–	EC
Survey	Production expectations, industry, EU	0	–	EC
Survey	Order books, industry, DE	0	–	EC
Survey	Production expectations, industry, DE	0	–	EC

Survey	Consumer confidence, DE	0	–	EC
Survey	Cons. major purchases over next 12 m., DE	0	–	EC
Survey	Cons. major purchases over next 12 m., EU	0	–	EC
Survey	Germany manuf. PMI, composite index	0	–	Bloomberg
Survey	Eurozone manuf. PMI, composite index	0	–	Bloomberg
Labour market	Wages and salaries	7	$\Delta \log$	SO SR
Trade balance	World trade	8	$\Delta \log$	CPB Netherlands

Note: For data availability reasons, PMI for years 2000-2001 backcasted using a regression with industrial confidence indicator as explanatory variable. Government bond yields for the year 2000 backcasted using a fixed interest differential against the 3m Bribor rate.