

Bayesian Averaging vs. Dynamic Factor Models for Forecasting Economic Aggregates with Tendency Survey Data

Piotr Bialowolski, Tomasz Kuszewski, and Bartosz Witkowski

Abstract

The article compares forecast quality from two atheoretical models. Neither method assumed a priori causality and forecasts were generated without additional assumptions about regressors. Tendency survey data was used within the Bayesian averaging of classical estimates (BACE) framework and dynamic factor models (DFM). Two methods for regressor selection were applied within the BACE framework: frequentist averaging (BA) and frequentist (BF) with a collinearity-corrected version of the latter (BFC). Since models yielded multiple forecasts for each period, an approach to combine them was implemented. Results were assessed using in- and out-of-sample prediction errors. Although results did not vary significantly, best performance was observed from Bayesian models adopting the frequentist approach. Forecast of the unemployment rate were generated with the highest precision, followed by rate of GDP growth and CPI. It can be concluded that although these methods are atheoretical, they provide reasonable forecast accuracy, no worse to that expected from structural models. A further advantage to this approach is that much of the forecast procedure can be automated and much influence from subjective decisions avoided.

JEL C10 C38 C83 E32 E37

Keywords Bayesian averaging of classical estimates; dynamic factor models; tendency survey data; forecasting

Authors

Piotr Bialowolski, ✉ Institute of Statistics and Demography, Warsaw School of Economics, Warsaw, Poland, piotr.bialowolski@sgh.waw.pl; piotr.bialowolski@gmail.com

Tomasz Kuszewski, Institute of Econometrics, Warsaw School of Economics, Poland

Bartosz Witkowski, Institute of Econometrics, Warsaw School of Economics, Poland

Citation Piotr Bialowolski, Tomasz Kuszewski, and Bartosz Witkowski (2015). Bayesian Averaging vs. Dynamic Factor Models for Forecasting Economic Aggregates with Tendency Survey Data. *Economics: The Open-Access, Open-Assessment E-Journal*, 9 (2015-31): 1—37. <http://dx.doi.org/10.5018/economics-ejournal.ja.2015-31>

1 Introduction

In the history of macroeconomic forecasting, two major trends led to two different approaches to economic modelling and forecasting. One group of models is based on inclusion of stylized facts from macroeconomic theory and thus causal effects are incorporated in modelling, while the other group of methods is atheoretical, based only on the observed properties of time series. Although inclusion of structural relations seems well justified, there are studies showing that accuracy of predictions obtained from such models is low (Kolasa et al., 2012; Rubaszek and Skrzypczyński, 2008). The second avenue, although not fundamentally better in terms of forecast quality, benefits from much less conceptual input during estimation. Hence, atheoretical methods are more easily implemented and accessible to a wider spectrum of potential users. However, this is not a path void of problems and investigation of the best methods for specific applications, like a single country forecasts, might be crucial to obtaining valid predictions. As the primary interest here was to provide quick and reliable forecasts with tendency survey data, the second path was the most appropriate choice.

The use of economic models without reference to economic theory for forecasting is by no means a new idea. The origins of this approach can be traced back to a brief comparison between seven structural models of the US economy and simple ARIMA forecasts (Cooper, 1972). The fundamental finding of this study was that forecasts from time series models were better than those from large scale structural models. Additionally, the effort associated with construction and testing such models was substantially less. Examples are either ARIMA or VAR models. It is relevant to recall the main points in favour of Sims' (1980) approach:

- (1) There is no a-priori distinction between exogeneous and endogeneous variables, i.e., no causal relations built between categories describing behaviour of an economic system;
- (2) No constraint model parameter values are imposed, in particular, it is not assumed that certain parameters are zero, which leads to elimination of variables associated with these parameters in the final form of the model;
- (3) There is no search performed for an underlying economic theory, which could be primary with respect to the model.

Our approach to forecasting the main macroeconomic indicators relies strongly on (2) and (3). However, as focus was on the main macroeconomic indicators – the rate of gross domestic product growth (GDP), consumer price index (CPI) and rate of unemployment (UNE) – the above variables were treated as endogenous. The novelty of the approach described here, is the exploitation of data sets from tendency surveys which contain numerous time-series describing economic behaviour. To deal with the volume of information, an attempt was made to introduce data mining techniques to macroeconomic forecasting. In order to benefit from the information carried by tendency survey data, approaches based on Bayesian averaging and dynamic factors were proposed.

The article primarily aims to compare atheoretical approaches to macroeconomic forecasting. A series of atheoretical models were designed to forecast the three main macroeconomic indicators quarterly: GDP, UNE and CPI. Generally, the use of explanatory variables is recommended, together with lagged values of these indicators, as well as current and lagged balances for various tendency survey items and composite indicators which are based on them. Competing models were evaluated with respect to their in- and out-of-sample forecasting performance. Although arguments for the use of forecasting models with tendency survey data and Bayesian averaging of classical estimates have already been made (Białowolski et al., 2012, 2014a), here the innovation is a twofold analysis, comprising both the approach known as „frequentist” (applied in the previous papers), adopting Bayesian averaging for the purpose of *selection* of model variables and the approach known as “averaging”, in which the independent variables are not selected but the results are averaged from different model structures with all possible regressors. In addition, a large set of Polish tendency survey data was applied for the first time to the dynamic factor framework for forecast of the main macroeconomic variables. Three sets of forecasts were generated for comparison.

Following these objectives, the paper is structured as follows. The next section (Part 2) focuses on a brief overview of the methodology. In Part 3 we present the data used for estimating the econometric models and describe the statistical properties of the time series used. Part 4 describes the modelling outcomes and in Parts 5 and 6, the forecasts are evaluated.

2 Forecasting models

Bayesian averaging models. Throughout the study it was assumed that the main research thrust was towards explanation of GDP growth (GDP), rate of inflation (CPI) and rate of unemployment (UNE). Selection of variables was in accord with their importance in assessment of economic situation and also accessibility of items from tendency surveys. The natural solution is a three equation model, in which it is assumed that all three time series GDP_t , UNE_t and CPI_t are interrelated, in addition to autocorrelation of each of these variables. Conceptually, a starting point for such analysis would be a three equation VAR model described as:

$$\begin{aligned} GDP_t &= f_1(GDP_{t-1}, UNE_t, CPI_t, \mathbf{V}_1, \varepsilon_{1t}) \\ UNE_t &= f_2(UNE_{t-1}, GDP_t, CPI_t, \mathbf{V}_2, \varepsilon_{2t}) \\ CPI_t &= f_3(CPI_{t-1}, GDP_t, UNE_t, \mathbf{V}_3, \varepsilon_{3t}), \quad t = 1, \dots, T; k \in \{0,1,2,3,4\}, \end{aligned} \quad (1)$$

where \mathbf{V}_1 , \mathbf{V}_2 and \mathbf{V}_3 represent “any other specified explanatory variables”. These might mean: the first or any further lags of GDP, UNE and CPI respectively, as well as any exogenous variables, such as economic situation indicators. However, two different approaches in this paper (dynamic factor and Bayesian averaging) were adopted for the following reasons. The first priority was to obtain a model for short term forecasts of GDP, UNE and CPI. Thus, \mathbf{V}_1 , \mathbf{V}_2 and \mathbf{V}_3 might only contain endogenous variable lags and such variables whose values are known for the near future. It was established whether they comprised a set of coincident and leading indicators from tendency surveys which might serve as reasonable determinants of GDP, UNE and CPI. The advantage of tendency survey data is that indicators for a given quarter are available at the beginning of the quarter, which allows both nowcasting and forecasting economic variables. Furthermore, in the construction of leading indicators at the RIED (Research Institute for Economic Development at the Warsaw School of Economics), company and household expectations are surveyed regarding the economic situation in the near future. This makes it reasonable to use k -th lags of the business tendency indicators rather than their current values, making it possible to extend the forecast horizon further by an additional k periods (quarters). That unfortunately comes at a cost. The series of business tendency and consumer sentiment indicators described

in the next section began in 1996, and so only 68 quarterly observations are available until end, 2012, which renders the VAR approach inapplicable. Another further problem is selection of suitable indicators from the tendency surveys for the model. Firstly, the number of available indicators is high, even if attention is limited to those provided by RIED. Not only would that mean very few (or even negative, when additional lags for endogenous variables are considered) degrees of freedom in the specified model, but also multiple collinearity would present an issue. Naturally, just a few indicators for the V_1, V_2 and V_3 sets could be preselected, but that would be counter to the approach relevant in the article, i.e., atheoretical, and it is unrealistic to expect that a rationale for the choice of a given subset of all the available tendency survey indicators could be offered.

To overcome these problems the following Bayesian approach was proposed. The model (1) was first replaced with the following structure:

$$GDP_t = f_1(GDP_{t-1}, \mathbf{X}_{1,t-k}, \varepsilon_{1t}) \quad (2a)$$

$$UNE_t = f_2(\widehat{GDP}_t, UNE_{t-1}, \mathbf{X}_{2,t-k}, \varepsilon_{2t}) \quad (2b)$$

$$CPI_t = f_3(\widehat{GDP}_t, \widehat{UNE}_t, CPI_{t-1}, \mathbf{X}_{3,t-k}, \varepsilon_{3t}), \quad (2c)$$

$$t = 1, \dots, T; k \in \{0, 1, 2, 3, 4\},$$

where $\mathbf{X}_{i,t-k}, i = 1, 2, 3$, represents the set of tendency survey indicators from period $t-k$ influencing the GDP growth, the rate of unemployment and the rate of inflation respectively; $\varepsilon_{it}, i = 1, 2, 3$, represents the error terms for subsequent equations, $f_i, i = 1, 2, 3$, is a certain linear function, \widehat{GDP}_t is the theoretical rate of GDP growth obtained from the equation (2a) and \widehat{UNE}_t is the theoretical rate of unemployment obtained from the equation (2b). Estimating (1) on an equation-by-equation basis would be inadequate due to endogeneity of particular variables. To overcome endogeneity, 2SLS-type logic was used to replace given variables with their theoretical values making recursive estimation of the model (2) feasible, simply using the least squares estimator. The sequence of equations is based on previous findings (Białowolski et al., 2010).¹ The classical assumptions were

¹ Naturally one could order the dependent variables in (2a)-(2c) in six different ways {GDP, UNE, CPI}, {GDP, CPI, UNE}, {CPI, GDP, UNE}, {CPI, UNE, GDP}, {UNE, GDP, CPI}, {UNE, CPI, GDP}, yielding six different sets of recursive equations. However, as shown in Białowolski et al.

adopted to estimate subsequent equations with the use of OLS: in particular, the error term was regarded spherical.

The next issue was selection of the optimal $X_{i,t-k}, i = 1,2,3$ for a given k . Firstly, it is not clear which tendency survey indicator lags should be used to maximize the forecast quality, except that it seems obvious that those should not be lagged too far. For that reason the set of (2a)–(2c) was estimated separately for different k between 0 (current values of tendency survey indicators) up to their 4th lags, without mixing different lags in one equation. It would be tempting to use more lags for the same indicator in the same equations (say, 1st and 2nd lags of them in one model), but this would be problematic, owing to very strong autocorrelation in most indicator series and high resultant multicollinearity. The subsequent issue was which indicators should be selected for a particular vector $X_{i,t-k}, i = 1,2,3$. Clearly, the set of indicators that would serve best as determinants for unemployment need not be the same as those for CPI or rate of GDP growth, thus each of the X vectors should be selected individually. The economic rationale in this case would be not only subjective but also the selection would require adopting a general-to-specific approach, which has been widely criticised (see, e.g., Ulaşan, 2012). Hence, the averaging approach was adopted using the Bayesian model for the purpose, which when OLS is used as estimator, degenerates to Bayesian averaging of classical estimates.

The technical details of Bayesian model averaging can be found in numerous papers, such as the milestone article of Sala-i-Martin et al. (2004) or Próchniak and Witkowski (2013), thus only a brief description of the procedure needs to be mentioned here.

Suppose we have $H = \{V_1, V_2, \dots, V_C\}$ as a set of C variables intended for inclusion in the estimator (the lagged endogeneous variables in our case). Further, let $X = \{Z_1, Z_2, \dots, Z_K\}$ be a set of K variables which are the tendency survey indicators, considered potential regressors in the equation (with $Z_k \notin H$ for $k=1, \dots, K$). There are exactly 2^K different linear regressions with a presumed dependent variable, all elements of H , as well as one of the 2^K possible subsets of X (including the empty set) as regressors. In the case of a low K , all the possible 2^K models denoted as M_1, \dots, M_J (J being thus equal to 2^K) are estimated. Also the

(2010), this way of ordering provided the set of equations that allowed for obtaining the most accurate forecasts in the past.

subset of X used in $M_j, j=1, \dots, J$ is denoted as X_j and the number of elements in M_j equals K_j . However, the number of models to be estimated in such a case increases dramatically with K . Thus usually instead of estimating all 2^K models, a large number of X_j 's are drawn and models based only on the selected subsets of X are estimated and further analysed ($J \ll 2^K$). At this point our approach is similar to the Stochastic Search Variable Selection (SSVS; George and McCulloch, 1993). Each of the subsequent estimated M_j 's can be viewed as vectors of K dummies, which indicate whether a given $Z_k \in M_j$. However, the classical SSVS was invented especially for the case when K is big, possibly very big and greatly exceeds the number of observations. This means that an efficient sampling design is needed to make the process efficient: Gibbs sampler is used and at the end of each iteration, the marginal effect of each variable's inclusion/exclusion is examined in order to decide the next iteration. This, however, was not essential in this case where K is not that large: the estimation process remains feasible if the subsequent M_j are drawn independently until convergence of parameter estimates is achieved.

In the next step of the model averaging procedure, an assumption regarding the prior distribution is needed. Which Z_k 's have real influence on the regressor in a given equation is not known. A popular and reasonable choice (usually called *binomial priors*) is to assume a certain value of \bar{k} , as the number of Z_k s in the true model.² Further, assuming independence of the potential regressors, the prior probability for each Z_k equals $\frac{\bar{k}}{K}$ and the prior probability for model M_j is

$$P(M_j) = \left(\frac{\bar{k}}{K}\right)^{K_j} \left(1 - \frac{\bar{k}}{K}\right)^{K-K_j} \quad (3)$$

Let D be the dataset used. The posterior probabilities of particular models $P(M_j | D)$, that is the probabilities of relevance for each M_j , can be written as:

$$P(M_j | D) = \frac{P(M_j)P(D | M_j)}{\sum_{i=1}^J P(M_i)P(D | M_i)}, \quad (4)$$

² Although for different \bar{k} the results need not be the same, we performed robustness checks for its values differing between 25% and 50% of the considered regressors, concluding no particular differences in our case.

which in the case of a linear model can be calculated as

$$P(M_j | D) = \frac{P(M_j) n^{-(K_j+C)/2} \text{SSE}_j^{-n/2}}{\sum_{i=1}^J P(M_i) n^{-(K_i+C)/2} \text{SSE}_i^{-n/2}}, \quad (5)$$

where n is the total number of observations in the dataset D while SSE_j is the sum of squared residuals of M_j . These can be viewed as the corrected probabilities of particular sets of potential regressors being the relevant, “best” ones.³

Further steps depend on the approach adopted. There are two types of Bayesian-averaging, which can be found in the literature: the “frequentist” and the “averaging” procedure (Moral-Benito, 2013). In this study, with regards to Bayesian averaging, three types of approaches were analysed: the averaging approach (BA), the frequentist approach (BF) and the frequentist approach with the control of collinearity (BFC). Adopting the BA approach, we next find the estimates of $\beta_{z_1}, \dots, \beta_{z_k}, \beta_{v_1}, \dots, \beta_{v_c}$ parameters treating the posterior probabilities (5) as weights. Let $\hat{\beta}_{r,j}$ be the estimator of a parameter in model M_j , let $\hat{\beta}_r$ be the ‘final’ estimator of parameter r , being the result of the total BA process. Let us denote their variances as $\text{Var}(\hat{\beta}_{r,j})$ and $\text{Var}(\hat{\beta}_r)$ respectively. Then

$$\hat{\beta}_r = \sum_{j=1}^J P(M_j | D) \hat{\beta}_{r,j}, \quad (6)$$

$$\text{Var}(\hat{\beta}_r) = \sum_{j=1}^J P(M_j | D) \cdot \text{Var}(\hat{\beta}_{r,j}) + \sum_{j=1}^J P(M_j | D) \cdot (\hat{\beta}_{r,j} - \hat{\beta}_r)^2. \quad (7)$$

³ There are also many alternative approaches based on selection of the final variable set with application of MCMC algorithm. For a review see O’Hara and Sillanpää (2009). Although these methods can be considered superior in terms of providing a unique solution, we opted for a methodology that allowed us to keep under control the issues of multi-collinearity characteristic for time-series analysis and tendency survey data.

The above means, that the parameter estimate is found for each potential regressor and each is then used for both inference and forecasting. In the BF approach we additionally define $P(Z_k | D)$ as the posterior probability of relevance for a given Z_k :

$$P(Z_k | D) = \sum_{j:Z_k \in X_j} P(M_j | D) \quad (8)$$

Using either the posterior probabilities for particular regressors or their pseudo- t statistics based on estimates of parameters and their variances (6,7), a subset of relevant X_j 's is selected (using the rule that for the selected X_j 's, the posterior probability should be no lower than the prior probability or that the variables should be statistically significant on the basis of the pseudo- t test: in this paper we use the pseudo- t statistic) and then the equation is selected only with its independent variables from the H set and the "relevant" X_j 's. However, in the latter procedure it might happen, that the selected regressors prove collinear. The BFC approach is therefore additionally proposed. In this case, after selecting the set of variables on the basis of their posterior probabilities, variance inflation factors were checked and the regressors with highest VIFs were eliminated recursively until all VIFs were acceptable (the usual $VIF < 10$ rule was adopted for this purpose). There is a risk of eliminating some relevant variables this way. However, if more variables are relevant and they are correlated with one another, they do not individually hold much additional information, so even despite the relevance of each, it might be wise to limit the set of regressors to those which are non-collinear.

Considering that 5 different sets of lags of $X_{i,t-k}, i = 1,2,3$ were considered ($k = 0,1,2,3,4$) and three above described approaches (averaging, frequentist, frequentist with collinearity correction) were tested, a total of 15 model structures were found. For every k and approach, firstly the equation (2a) was estimated and the theoretical values of GDP_t were found. When the case used the frequentist approach, those were the theoretical values of GDP from a single equation with a "Bayesian-selected" set of tendency survey indicators and the lagged GDP (having additionally eliminated the statistically collinear indicators in the collinearity corrected frequentist approach). For the averaging approach, averaged parameter estimates for each regressor were found from all the models M_j according to (6). Then the theoretical GDP was found as a linear predictor with the use of all the

considered regressors, i.e. $\widehat{GDP}_t = B_t \hat{\beta}_r$, where $B_t = [GDP_{t-1} \mathbf{X}_{1,t-k}]$. Then the process was repeated for equation (2b), except that the theoretical GDP from (2a) was used as one of the regressors (for each of the three approaches, theoretical GDP was obtained using the same approach applied to equation 2a). Finally, the same process was applied to equation (2c), while theoretical GDP from (2a) and theoretical unemployment rate from (2b) were used as independent variables. In all the Bayesian averaging models, only the prognostic variables from the tendency survey time series were used. Due not only to computational complexity of those methods but also to the research question oriented on forecasting, when adopting the Bayesian approach it was decided to omit indicators describing the current state of economic affairs or assessing the current climate. Adopting such an approach, it was possible to reduce significantly computations required to obtain results.⁴

Due to the considerable number of estimates generated by the Bayesian averaging procedure, it was decided to present only the set of regressors from the sets of X in equations (2a)-(2c). In the BA method, following the philosophy of this method, in each the three equations and for each lag k , the set of regressors from the tendency surveys was the same and comprised the following indicators (Appendix, Table A3). In the frequentist approach (BF, BFC) the set of regressors differed in models with collinearity correction and without it (Appendix, Table A4).

Analysis of explanatory variable patterns in the equations for macroeconomic variables enabled the following conclusions:

- The cases with exactly the same the set of indicators for models with and without collinearity correction imply that collinearity was not observed.
- The set of regressors depends on the lag (k). In the equations for GDP and CPI similarities are observed with in the sets: $\{k=0\}$, $\{k=1, k=2\}$, $\{k=3, k=4\}$, in the equations for UNE the sets are: $\{k=0, k=1\}$, $\{k=2, k=3, k=4\}$.
- A significant role is played by the regressors from consumer tendency surveys (CSO and RIED).

⁴ Detailed description of results achieved with the Bayesian approach can be found in (Białowolski et al., 2014b).

- The most frequently occurring indicators (except for the equation on GDP) are those from the Bureau of Investment and Economic Cycles - `biec_xxx`.

Dynamic factor models. Application of dynamic factor models to forecasting macroeconomic time series has been extensively developed in the literature (Baranowski et al., 2010; Boivin and Ng, 2006; Reijer, 2012; Stock and Watson, 2002). Nevertheless, with minor exceptions it has rarely focused on defining the dynamic factors with tendency survey data (Frøle et al., 2010; Hansson et al., 2005; Kaufmann and Scheufele, 2013). However, it should be underlined that dynamic factor models have significant advantages over other approaches to modelling. Breitung and Eickmeier (2006) enumerated advantages of the dynamic factor approach, summarized as follows: (1) Factor models can cope with many variables without excessive reduction of degrees of freedom, often the case when using many input variables to regression based modelling;⁵ (2) In factor models idiosyncratic movement of specific variables, which may include measurement error and local shocks, can be eliminated; (3) Dynamic factor models allow modellers to remain agnostic about the structure of the economy and without reliance on various assumptions, often necessitated by structural models.

With regards to forecasting, a particular advantage of dynamic factor models is the elimination of noise from data. Hansson et al. (2005) claim that the idiosyncratic processes present in different sectors are probably not relevant to general economic processes. Eliminating them with a factor approach might be crucial, when the focus of analysis is on macroeconomic aggregates, as in analysis described here. Dynamic factor models proved especially useful (see Point 3 above) as their structure and implied strategy matched initial assumptions, with limited influence from modellers impinging on the forecasting process.

Dynamic factor models also have certain drawbacks that should be considered. A disadvantage of common factor models is that factors may partially or completely lack clear interpretation. As a result, Stock and Watson (2002) suggested that they should be interpreted as diffusion indices oriented on assessment of average economic activity. Naturally, there are also caveats associated with a number of indicators. A larger number of indicators is not always

⁵ Time series models usually contain no more than 10 time series (Boivin and Ng, 2006; Stock and Watson, 2002). Even our approach based on Bayesian Averaging was constructed in such a way that the optimal number of time series in an equation should be around 6.

desirable even in the dynamic factor specification. Boivin and Ng (2006) showed that adding a series highly correlated with another might reduce rather than improve efficiency of factor estimates. Otherwise, adding a ‘noisy’ time series, that shares little common variance with other series also reduces the efficiency of factor estimates, since the average common component is reduced. So, the goal here, in establishing common factors was to pick a considerable number of time series from tendency surveys in the data set but at the same time eliminate series only contributing noise to the final factor solutions.

Regardless of the character of time series used, the structure of the dynamic factor model is similar. The starting point for analysis is an approximate factor model with K factors, taking the form:

$$\mathbf{X}_t = \mathbf{A}\mathbf{F}_t + \boldsymbol{\varepsilon}_t, \quad (9)$$

where \mathbf{X}_t represents $N \times 1$ vector of consumer and business tendency survey indicators (also composite indicators used in the analysis) measured at a given time t , \mathbf{A} is a matrix of factor loadings of dimension $N \times K$, \mathbf{F}_t is the $K \times 1$ vector of period specific factor loadings, $\boldsymbol{\varepsilon}_t$ is an $N \times 1$ vector of measurement errors over a given period.

Following Stock and Watson (2002), the number of factors here was determined according to the loadings from using the simple principal component approach.⁶ Additionally, it was assumed that the number of factors was determined from the standard Cattell criterion. In order to eliminate variables which have very low factor loadings, the assumption from other factor models was adopted, that loadings needed to be salient, which was assumed to apply to values over 0.5. Brown (2006) suggested a range between 0.4 and 0.6 for factor models based on individual data, however here it was assumed that the half interval was appropriate for dynamic factors.⁷ A drawback from dealing only with static

⁶ Naturally, for extraction of the common factors, a different factor analytical approach can be used, like exploratory factor analysis. Nevertheless, differences in the results (factor loadings) between various factor analytical approaches are usually very small and thus this issue was not subject to deep analysis.

⁷ A sensitivity check of the final results with 0.4 and 0.6 thresholds was also performed. The results were similar to the baseline scenario (average differences between forecasts did not exceed 0.1 percentage point even for the longest horizon). One can also set the threshold at a much lower level.

factors, is that the dynamic structure, which is likely to exist between the factors, might not be reflected. In order to account for these possible dynamics, a dynamic component was introduced, based on the static factors obtained. The dynamic factor model is an extended version of the static model, where the factors are assumed to follow the dynamic, autoregressive process:

$$F_t = \Phi(L)F_{t-1} + \mu_t, \quad (10)$$

where $\Phi(L)$ is a vector of lag polynomials describing the autoregressive structure of the data generating process for factors and μ_t describes the error. In our empirical approach, we assessed models with a lag polynomial in the form: 1, L, L², L³ and 1+L³, so, lags equal to 1,2,3,4 and 1 and 4 simultaneously were of interest. Selection of the appropriate lag was based on the Schwarz Information Criterion. The final step of the analysis, oriented on forecasting with dynamic factor models, was the inclusion of dynamic factors into the process. A standard specification for a model with dynamic factors used as a forecasting tool can be presented by the following system of equations (see Stock and Watson, 2002; Baranowski et al., 2010)

$$y_t = \alpha + \sum_{m=1}^L \beta_m y_{t-m} + \sum_{n=0}^L \gamma_n F_{t-n} + \varepsilon_t, \quad (11)$$

where y_t represents a vector of macroeconomic variables of interest, α stands for a vector of constants, L is the number of lags included in the analysis, β_m is a vector of autoregressive coefficients for variables of interest lagged by m periods and γ_n is a vector of coefficients for dynamic factors lagged by n periods.

In this case due to the intention to include interrelations between the current level of indicators, a slightly modified approach was taken. In previous studies the established order, according to which macroeconomic variables should be related to each other was defined by equations (2a–2c). Inclusion of the interrelations between the macroeconomic variables resulted in a slightly modified framework with dynamic factors used for the forecasting purposes. Having $y_t = [GDP_t \text{ } UNE_t \text{ } CPI_t]^T$ but also additional assumptions that only one lag of the variable of interest is included in the equation for this variable and that dynamic

However, such procedure is mostly implied for analyses dealing with micro level data, when a check that data fits the model is more important (see, e.g., Białowolski and Weziak-Białowolska, 2014).

factor estimates are taken only for a single quarter depending on the chosen lag (five possibilities of lags were checked $k = 0,1,2,3,4$), the final model can be presented by the following system:

$$\begin{aligned} GDP_t &= \alpha_{GDP} + \beta_{GDP}GDP_{t-1} + \sum_{i=1}^N \gamma_{i,GDP,k} F_{i,t-k} + \varepsilon_{t,GDP,k} \\ UNE_t &= \alpha_{UNE} + \kappa_{UNE} \widehat{GDP}_t + \beta_{UNE} UNE_{t-1} + \sum_{i=1}^N \gamma_{i,UNE,k} F_{i,t-k} + \varepsilon_{t,UNE,k} \\ CPI_t &= \alpha_{CPI} + \kappa_{CPI} \widehat{GDP}_t + \lambda_{CPI} \widehat{UNE}_t + \beta_{CPI} CPI_{t-1} + \sum_{i=1}^N \gamma_{i,CPI,k} F_{i,t-k} + \varepsilon_{t,CPI,k} \end{aligned} \quad (12)$$

In the final specification, in the second equation (for UNE) the estimated value of GDP for period t is included as an exogenous variable, while in the third equation (for CPI) both estimates of GDP and UNE are included as exogenous variables. In addition to this, all N dynamic factors are present in all equations (see Appendix, Table A5). Thus, although the variable selection procedure is significantly different, the modelling strategy implemented in the dynamic factor framework shares the final structure of forecasting models with Bayesian approaches, which serve as a tool for generating the final forecasts.

3 Data – sources and preparation

In order to build forecasting models, quarterly data covering the years from 1996 to 2014 were collected. The data on the gross domestic product (GDP), the consumer price index (CPI) and the unemployment rate (UNE) came from Poland's Central Statistical Office (CSO). The unemployment rate was set according to the Labour Force Survey. GDP, CPI and UNE served in our models as endogenous variables. The set of indicators was extended not only with time series on individual consumption, investment outlays, export and import but also value added in 16 sectors of the economy.

In addition to the lagged endogenous variables and data from national accounts, tendency survey data are assumed to play the role of regressors in the econometric models that were designed, either in their original form or as variables explained by the presence of common factors. Tendency survey data is usually published in the form of monthly statistics. In line with standard practice, business survey data for the first month of each quarter, i.e., January, April, July and

October, are considered indicators for their respective quarters. The database applied in the procedure comprises a time series from the Research Institute for Economic Development (RIED) at the Warsaw School of Economics (WSE), on sentiment in the manufacturing industry, trade and construction and households. Data published by the Centre for European Economic Research (ZEW), the Leibniz Institute for Economic Research at the University of Munich (Ifo Institute), Bureau for Investments and Economic Cycles (BIEC), and the Purchasing Managers' Index (PMI) for Polish industry, were also collected and subsequently applied in the analysis. In addition to this, data on consumer confidence from the Central Statistical Office and IPSOS group were included. The symbols adopted for the variables in the estimated models are presented in the Appendix.

Similar to most empirical studies, data generating processes were verified with respect to stationarity. Most research provides verification of stationarity with respect to the mean,⁸ which is usually accounted for by differencing the time series. In this case, stationarity was checked with ADF and KPSS tests (Kwiatkowski, Phillips, Schmidt, and Schin, 1992) to study order of integration. No time series with an order of integration higher than 1 were identified in the database. Nevertheless, the analysis showed that it could be assumed that the time series for responses to business survey questions targeting the industrial sector were stationary $I(0)$, while the series for responses to business survey questions targeting households were integrated $I(1)$. The remaining regressor time series appeared to be stationary. This explains the decision made, against differentiating the values of the series $I(1)$; instead statistical properties of the residual series from the estimated models were studied. Stationarity of regressand time series was investigated with KPSS. The time series for GDP was stationary, but CPI and UNE are integrated at degree 1 ($d=1$).

Discussion regarding the seasonality of time series is ever present in the literature (see, e.g., Clements and Hendry, 2011). The voices of those favouring of de-seasoning in economic modelling number more or less equal to those against it. However, seasonality treatment of time series was omitted in this analysis, since

⁸ Stationarity with respect to variance is rarely subject to verification. Lack of stationarity with respect to variance is usually accounted for by taking the logarithm of the time series. However, such a procedure appeared unnecessary in this case.

the results presented in Białowolski et al. (2014a) indicated only marginal influence for both deterministic and stochastic specifications of a seasonal factor. This follows a common econometric finding that with either version of seasonality (deterministic or stochastic), due to the fact that different patterns of seasonality are present among regressors, it is hard to predict the influence of seasonality on parameter estimates and, more importantly, on forecasts (see, e.g., Mycielski, 2010 for more information).

In the literature, arguments can be found to substantiate that de-seasonized time-series are in fact obtained via estimation and due to this some of the information content of time series subject to de-seasoning is lost (see e.g. Bloem et al., 2001). It has been also pointed out that seasonality correction should be performed when the same months or quarters are compared for different years in an analysis of a single times-series, while seasonal correction is less justified when the time-series data serve for modelling economic processes (Manski, 2014). As an example, in the macro-econometric model for the Polish economy WK2009 (Welfe, 2013) based on quarterly data, only non-seasonally adjusted data were used.

The influence of de-seasoning a time-series on quality of estimates and the testing of autoregressive models was assessed by Hecq (1998). He obtained strong support against application of seasonal treatment to time-series data. However, if time-series are to be used in applications other than econometric modelling, seasonal treatment might be better justified (Baranowski et al., 2010). Consequently, it was decided to use raw time series in all models.

4 Fitting forecasting models to the data

Fit of the forecast models can be measured in two ways. The first involves the analysis of signs of the differences between the empirical and theoretical regressands in successive periods. Comparison of the signs allows judgement of the reaction of the models to the change in direction of trends in the macroeconomic indicators. The second possibility to verify the quality of fit is to use of one of the standard measures of ex-post errors used in forecast analysis. This measure is the mean absolute percentage error (MAPE), the mean value of error expressed as a percentage of the true value of the analysed variable. MAPE

allows valid comparison of fit and forecast accuracy independently from units of the regressand.⁹

The conclusions from comparison of the accuracy of the model fit in the period from the 1st quarter, 1996 to the 4th quarter, 2012 (Table 1) are the following:

- Considering the number of coinciding signs of the differences of the empirical and the theoretical values of the regressands, no method is clearly superior to any other. This infers that all models correctly identified direction of change in short term trends in approximately the same percentage of cases (roughly exceeding 2/3 of all attempts).
- The number of coinciding signs of the differences for all the models and all the macroeconomic indicators decreased in line with increasing lag of variables from tendency surveys.
- The analysis of the MAPE allows ordering of the forecast models, starting with the best fit. The Bayesian frequentist approach is characterized by lower MAPEs than models identified by dynamic factor analysis, while models identified using classical Bayesian model averaging function poorest in this respect.
- MAPEs increase with lag times of the variables from tendency surveys.
- Whichever algorithm was used, for each model, it is the equation which explains the UNE variable fitting the data best, followed by the CPI equation, while the GDP equation proves to be the least accurate fit. Considering that the order of equations in the model is the same in each case, it can be seen clearly, that errors estimating GDP are not accumulated with the errors related to estimation of the other two variables.

In the next step the forecast errors in the period from the 1st quarter, 2013 to the 2nd quarter, 2014 were considered. However, the model identified by averaging the classical Bayesian model was eliminated from further analysis as the worst fit.

⁹ In the literature, there has been long ongoing discussion on the sense of different measures of accuracy. Some believe that MAPE is not appropriate. Hyndman and Koehler (2006) proposed measures that stem from MAPE but differ from it in their construction. However, these add no additional interpretational possibilities to this paper.

Table 1: Quality of the estimates of forecasting models measured with the use of the regressands' sign difference and MAPE for the 1996q1–2012q4 period

| Regressand | DFM models | | | | | BA models | | | | |
|------------|---|---------|---------|---------|---------|------------|---------|---------|---------|---------|
| | $k = 0$ | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 0$ | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ |
| | Number of coinciding signs of the differences of empirical and theoretical values | | | | | | | | | |
| GDP | 47 | 43 | 44 | 44 | 43 | 45 | 44 | 42 | 39 | 42 |
| UNE | 37 | 39 | 40 | 40 | 36 | 45 | 45 | 42 | 45 | 49 |
| CPI | 48 | 47 | 44 | 45 | 44 | 45 | 46 | 47 | 41 | 45 |
| | MAPE | | | | | | | | | |
| GDP | 0.31 | 0.31 | 0.31 | 0.34 | 0.34 | 0.62 | 0.39 | 0.37 | 0.50 | 0.38 |
| UNE | 0.05 | 0.05 | 0.05 | 0.05 | 0.06 | 0.11 | 0.12 | 0.11 | 0.12 | 0.06 |
| CPI | 0.15 | 0.17 | 0.20 | 0.21 | 0.21 | 0.87 | 0.25 | 0.26 | 0.40 | 0.40 |
| Average | 0.17 | 0.18 | 0.19 | 0.20 | 0.20 | 0.53 | 0.25 | 0.24 | 0.34 | 0.28 |
| | BF models | | | | | BFC models | | | | |
| | $k = 0$ | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 0$ | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ |
| | The number of coinciding signs of the differences of empirical and theoretical values | | | | | | | | | |
| GDP | 44 | 45 | 41 | 42 | 41 | 44 | 43 | 41 | 42 | 39 |
| UNE | 45 | 47 | 43 | 44 | 45 | 44 | 47 | 40 | 40 | 44 |
| CPI | 49 | 44 | 46 | 44 | 45 | 47 | 43 | 48 | 43 | 40 |
| | MAPE | | | | | | | | | |
| GDP | 0.24 | 0.24 | 0.28 | 0.25 | 0.28 | 0.24 | 0.24 | 0.28 | 0.27 | 0.27 |
| UNE | 0.04 | 0.04 | 0.04 | 0.04 | 0.03 | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 |
| CPI | 0.16 | 0.15 | 0.17 | 0.17 | 0.17 | 0.15 | 0.15 | 0.18 | 0.18 | 0.19 |
| Average | 0.14 | 0.15 | 0.16 | 0.15 | 0.16 | 0.14 | 0.15 | 0.17 | 0.16 | 0.16 |

Source: own estimates. Remark: For every k which represents the delay of lag a different number of differences were compared; 66 for $k = 0$, 66 for $k = 1$, 65 for $k = 2$, 64 for $k = 3$ and 63 for $k = 4$.

5 Forecasting

In the discussion on the construction of the forecasting model it has been already emphasized that more than one forecast of each macroeconomic indicator can be

made as a virtue of the differing lags of tendency survey variables to describe business climate. The set of forecasts for each indicator obtained using the same empirical observations can be identified as a portion of forecasts.¹⁰ With survey data accessible in the first month of a quarter nowcasting, i.e., forecasting with zero lag ($k=0$ delay), is feasible. Further, models used to forecast with lags $k=1,2,3,4$ have also been prepared. An example of portions of forecasts obtained by the frequentist approach with collinearity correction (Table 2) included 15 values for each regressand: one single value for $k=0$ and five for $k=4$. Table 2 presents the forecasts generated with the use of data on GDP, UNE and CPI ending in the 1st and 2nd quarter of 2014.

The analysis of ex post forecast accuracy, based on the separate portions of forecasts, answers the question about the stability of the forecast process as successive observations are piece-wise added to the regressor time series. Models were identified in the sense of their general functional form (that is the set of regressors) with the use of time series of independent variables ranging from the 1st quarter, 1996 to the 4th quarter, 2012. Estimates for the structural parameters changed (re-estimated on the time-constant regressor set) with each new observation added to the series. An example of ex post forecast accuracy based on the portion of forecasts using time series ranging from the 4th quarter, 2012 to the 1st quarter, 2014 was presented in Table 3.

By extending the time series by additional quarters, values of MAPE can be compared. It is worth noting that these forecasts were obtained as combinations of portions of forecasts of varying size, ranging from 5 to 15 values. The unemployment rate was found to be forecast with greatest precision, followed by the rate of GDP growth and the CPI.

¹⁰ Empirical values from tendency surveys which correspond to the 1st quarter were from January, those for the 2nd quarter from April, those for the 3rd quarter – from July while those for the 4th quarter were from October.

Table 2: Forecasts from BFC model

| k | Last period of data 2014q1. Forecasts for: | | | | | Last period of data 2014q2. Forecasts for: | | | | |
|---|--|------------|------------|------------|------------|--|------------|------------|------------|------------|
| | 2014q 2 | 2014q 3 | 2014q 4 | 2015q 1 | 2015q 2 | 2014q 3 | 2014q 4 | 2015q 1 | 2015q 2 | 2015q 3 |
| | GDP | | | | | | | | | |
| 0 | 3.8 | | | | | 3.2 | | | | |
| 1 | 3.8 | 4.5 | | | | 4.1 | 4.2 | | | |
| 2 | 3.8 | 4.1 | 4.4 | | | 3.6 | 3.9 | 4.2 | | |
| 3 | 3.5 | 3.6 | 3.6 | 3.2 | | 3.4 | 3.4 | 3.0 | 3.5 | |
| 4 | 3.1 | 2.8 | 2.4 | 2.2 | 2.1 | 3.0 | 2.5 | 2.4 | 2.2 | 2.2 |
| | UNE | | | | | | | | | |
| 0 | 9.6 | | | | | 9.1 | | | | |
| 1 | 10.4 | 10.0 | | | | 8.7 | 8.4 | | | |
| 2 | 10.8 | 10.6 | 10.9 | | | 9.1 | 9.3 | 8.4 | | |
| 3 | 10.5 | 10.1 | 9.6 | 9.7 | | 8.8 | 8.5 | 8.7 | 7.8 | |
| 4 | 9.8 | 8.8 | 7.8 | 7.5 | 5.8 | 7.9 | 6.8 | 6.3 | 4.5 | 4.3 |
| | CPI | | | | | | | | | |
| 0 | 0.6 | | | | | 0.6 | | | | |
| 1 | 0.6 | 1.2 | | | | 0.8 | 0.6 | | | |
| 2 | 0.9 | 1.2 | 1.7 | | | 0.4 | 0.7 | 0.5 | | |
| 3 | 0.5 | 0.4 | 0.2 | -0.1 | | 0.1 | -0.3 | -0.6 | -0.6 | |
| 4 | 0.6 | 0.4 | 0.2 | -0.2 | -0.3 | 0.1 | -0.2 | -0.5 | -0.7 | -0.7 |

Source: own estimates.

6 Combined forecasts

Judging from the structure of the forecasts, a choice of many forecasts for a single macroeconomic indicator is clearly available in any quarter. However, if a single, point forecast were to be preferred, the task would be to “average” the forecasts obtained. In the situation described, the additional difficulty stems from the differing number of forecasts according to the number of preceding quarters used in the forecast. Further, it should be appreciated that forecasts may lose accuracy with increase in the lag between the forecast period and the last observation.

Table 3: The MAPE-measured accuracy of the portion of forecasts based on the gradually extended time series

| Regressand | Last period of data and number of forecasts in portion | | | | | |
|------------|--|--------|--------|--------|--------|--------|
| | 2012q4 | 2013q1 | 2013q2 | 2013q3 | 2013q4 | 2014q1 |
| | 15 | 15 | 14 | 12 | 9 | 5 |
| | DFM models | | | | | |
| GDP | 0.75 | 0.25 | 0.30 | 0.16 | 0.18 | 0.13 |
| UNE | 0.14 | 0.27 | 0.10 | 0.07 | 0.07 | 0.17 |
| CPI | 0.76 | 0.56 | 1.00 | 0.88 | 1.18 | 1.22 |
| Average | 0.55 | 0.36 | 0.47 | 0.37 | 0.48 | 0.51 |
| | BF models | | | | | |
| GDP | 0.52 | 0.37 | 0.31 | 0.19 | 0.18 | 0.13 |
| UNE | 0.11 | 0.19 | 0.11 | 0.07 | 0.08 | 0.13 |
| CPI | 0.90 | 1.27 | 1.56 | 1.38 | 0.93 | 1.00 |
| Average | 0.51 | 0.61 | 0.66 | 0.55 | 0.40 | 0.42 |
| | BFC models | | | | | |
| GDP | 0.51 | 0.31 | 0.30 | 0.16 | 0.18 | 0.11 |
| UNE | 0.09 | 0.19 | 0.11 | 0.07 | 0.07 | 0.12 |
| CPI | 0.81 | 0.73 | 1.00 | 0.88 | 1.18 | 1.06 |
| Average | 0.47 | 0.41 | 0.47 | 0.37 | 0.48 | 0.43 |

Source: own estimates.

The number of generated forecasts in relation to the length of forecasting horizon is illustrated by table 4. For the first time, a GDP forecast for the 1st quarter, 2014 was obtained from the model using last observed data from the 4th quarter, 2012, when the lag order was assumed as $k=4$. In the next step, when information regarding the 1st quarter 2013 was already available, two forecasts could be made (for $k=3$ and $k=4$). Finally, when the data on macroeconomic indicators up to the 4th quarter, 2013 were gathered, the 1st quarter 2014 forecast became available for all $k=0,1,2,3,4$. Consequently, having the information up to the 4th quarter, 2013, 15 forecasts could be obtained for five time points (quarters). The dispersion of the values and the number of calculated forecasts for the 1st quarter, 2014 is illustrated in Figure A1–A3 (Appendix). It can be seen that the shorter the lag for regressors from tendency surveys, the closer the forecasts of a given macroindicator are to reality.

In the process of aggregation of the forecasts with different forecasting horizons, weights are applied. These should be non-negative real numbers with

their sum equal to one. It is also assumed that the forecast made in period t for a given quarter is more important than the forecast at period $t-1$. Finally, it is assumed that the second derivative of a weight with respect to t is nonnegative. The last condition is driven by the assumption that the difference in importance between the information from time point t and information from point $t-1$ is at least as great as the difference in importance between the information available at $t-1$ and $t-2$. A family of weight functions satisfying this condition can be shown (Czerwiński and Guzik, 1980). The most popular are harmonic, linear and exponential weights (Table 4). The weights are usually described by a sequence of m observations ordered with respect to t ($t=1,2,\dots,m$) given the following formulas:

– harmonic weights $w_t^m = w_{t-1}^m + \frac{1}{m(m-t+1)}$, $t = 1, 2, \dots, m$; $w_0^m = 0$

– linear weights $w_t^m = \frac{2t}{m(m+1)}$, $t = 1, 2, \dots, m$

– exponential weights $w_t^m = \frac{(1-q)q^{m-1}}{1-q^m}$, $t = 1, 2, \dots, m$; $0 < q < 1$.

Growth of harmonic weight is proportional to the difference between m and t . Differences in the linear specification of weights are constant. Differences in weights obtained in the exponential weighting procedure increase with the growth of t . Exponential weights have an additional important property. By taking an adequate value for q , the declining importance of past observations can be accommodated.

A two-step procedure was used to obtain the combined forecast. First, the arithmetic mean of all the forecasts available for the given period was calculated. So, for example, all the 5 forecasts for the 1st quarter, 2014 were computed (using $k=0,1,2,3,4$ lags) and means found using values of regressors from up to the 4th quarter, 2013. The time between the forecast and the last observed empirical value was one quarter ($m=1$). Using the values of regressors up to the 3rd quarter, 2013, 4 forecasts for the 1st quarter, 2014 (with the use of $k=0,1,2,3$ lags) were computed and averaged. The time between the forecast and the last observed empirical value was two quarters ($m=2$). The same was repeated for survey data up to the 2nd

Table 4: The weights for different forecast lags

| Weights | Weights for the forecasts made a given number of quarters back | | | | |
|-------------|--|------------|------------|------------|-----------|
| | 5 quarters | 4 quarters | 3 quarters | 2 quarters | 1 quarter |
| Harmonic | 0.04 | 0.09 | 0.16 | 0.26 | 0.45 |
| Linear | 0.07 | 0.13 | 0.20 | 0.27 | 0.33 |
| Exponential | 0.00 | 0.01 | 0.03 | 0.16 | 0.80 |

Source: own estimates.

Table 5: Combined forecasts for quarters 2014q1 and 2014q2

| | | | | | | |
|--------------------------------|---------------|--------|--------|---------------|--------|--------|
| Last period of data | 2012q4 | 2013q1 | 2013q2 | 2013q3 | 2013q4 | 2014q1 |
| Number of forecasts for 2014q1 | 1 | 2 | 3 | 4 | 5 | |
| Average from forecasts GDP | 1.9 | 2.3 | 2.0 | 2.9 | 3.4 | |
| UNE | 11.5 | 11.8 | 11.5 | 10.5 | 10.1 | |
| CPI | 0.9 | 0.4 | -0.3 | 0.6 | 0.7 | |
| Number of forecasts for 2014q2 | | 1 | 2 | 3 | 4 | 5 |
| Average from forecasts GDP | | 3.3 | 3.0 | 3.5 | 4.2 | 3.6 |
| UNE | | 12.0 | 10.4 | 10.6 | 10.1 | 10.2 |
| CPI | | 0.8 | 0.4 | 0.9 | 1.0 | 0.6 |
| Combined forecasts for | 2014q1 | | | 2014q2 | | |
| Weights | Harmonic | Linear | Exp | Harmonic | Linear | Exp |
| GDP | 2.9 | 2.7 | 3.3 | 3.7 | 3.7 | 3.7 |
| UNE | 10.7 | 10.8 | 10.3 | 10.3 | 10.4 | 10.2 |
| CPI | 0.5 | 0.5 | 0.7 | 0.7 | 0.8 | 0.7 |
| Real GDP | 3.4 | | | 3.5 | | |
| Real UNE | 10.6 | | | 9.1 | | |
| Real CPI | 0.6 | | | 0.3 | | |

Source: own estimates.

quarter, 2013, the 1st quarter, 2013 and the 4th quarter, 2012, with 3, 2, and 1 obtained forecasts. Then, in the second stage of the averaging process, three types of weights for various forecast lags ($m=1,2,3,4,5$) were used and combined forecasts were calculated (Table 5).

The reason for providing the 1st quarter, 2014 and the 2nd quarter, 2014 forecasts is that the full set of 15 forecast values was available for these quarters. No particular method of averaging which uses just this information can be recommended, since the number of forecasts was insufficient to justify drawing any definite conclusions.¹¹

7 Concluding remarks

In this study, a prognostic model was constructed using three methods for the three key macroeconomic indicators: GDP growth, the unemployment rate and the consumer price index. Two methods applied variations of Bayesian averaging methods (“averaging” and “frequentist”) and the third adopted the dynamic factor approach. All exploited the set of indicators from tendency surveys. The collection procedure for business and consumer sentiment indicators and approach to use of lagged values of tendency survey data as regressors permitted forecast generation without additional assumptions regarding outcomes. The method eliminated all subjective assumptions concerning economic processes and typically made by forecasters. Arguably, forecaster intuition is substituted by aggregate intuition manifest in business and consumer tendency survey data itself.

Forecasts from the Bayesian approaches confronted those yielded by the dynamic factor model and results demonstrated the best performance of the “frequentist” approach, characterized by the lowest, mean in-sample and out-of-sample absolute percentage errors. Differences in forecast error between the Bayesian and dynamic factor models were however very small, suggesting that

¹¹ The computation described above should be considered as an illustration if the forecasting practice is continued. Białowolski et al. (2014b) provided a more detailed overview of forecasting performance for the chosen set of models between the 1st quarter, 2013 and the 1st quarter, 2014. They compared their results with simple autoregressive estimates and with forecasts made by the Economic Institute of the National Bank of Poland in addition to forecasts prepared by the Institute of Market Economy Research.

both approaches had similar forecasting efficiency. This is confirmed by the very minor differences in aggregate forecasts for the 1st and the 2nd quarter of 2014.

An important innovation offered by our approach is that it lends itself to being largely automated, significantly reducing the influence of subjective decisions otherwise needed in forecasting. Encouragingly, the forecast methods successfully combined statistical and econometric methodologies with the data mining approach.

Acknowledgements The authors would like to thank the National Bank of Poland for financing this work within its scientific grant framework – grant „Prognozowanie podstawowych wskaźników makroekonomicznych z wykorzystaniem uśredniania bayesowskiego oraz modeli czynnikowych w oparciu o dane z testów koniunktury”. Authors also thank two anonymous reviewers, participants of seminar in the National Bank of Poland and Patrick Fox for their most helpful comments. Remaining errors are ours.

References

- Baranowski, P., Leszczyńska, A., and Szafranski, G. (2010). Krótkookresowe prognozowanie inflacji z użyciem modeli czynnikowych. *Bank & Credit* 41(4): 23–44. http://www.bankikredyt.nbp.pl/content/2010/04/bik_04_2010_02_art.pdf
- Białowolski, P., and Weziak-Białowolska, D. (2014). External factors affecting investment decisions of companies. *Economics. The Open-Access, Open-Assessment E-Journal* 8(2014-11): 1–22. <http://www.economics-ejournal.org/economics/journalarticles/2014-11>
- Białowolski, P., Kuszewski, T., and Witkowski, B. (2010). Business survey data in forecasting macroeconomic indicators with combined forecasts. *Contemporary Economics* 4(4): 41–58. <http://we.vizja.pl/en/issues/volume/4/issue/4#art178>
- Białowolski, P., Kuszewski, T., and Witkowski, B. (2012). Macroeconomic forecasts in models with Bayesian averaging of classical estimates. *Contemporary Economics* 6(1): 60–69. <http://we.vizja.pl/en/download-pdf/volume/6/issue/1/id/232>
- Białowolski, P., Kuszewski, T., and Witkowski, B. (2014a). Bayesian averaging of classical estimates in forecasting macroeconomic indicators with application of business survey data. *Empirica* 41(1): 53–68. <http://doi.org/10.1007/s10663-013-9227-x>

- Białowolski, P., Kuszewski, T., and Witkowski, B. (2014b). *Dynamic factor models & Bayesian averaging of classical estimates in forecasting macroeconomic indicators with application of survey data* (No. 191). Warsaw.
http://www.nbp.pl/publikacje/materialy_i_studia/191_en.pdf
- Bloem, A.M., Dippelsman, R.J., and Maehle, N.O. (2001). *Quarterly national accounts manual - concepts, data sources and compilation*. Washington D.C.: International Monetary Fund.
<https://www.imf.org/external/pubs/ft/qna/2000/Textbook/ch1.pdf>
- Boivin, J., and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics* 132(1): 169–194. <http://doi.org/10.1016/j.jeconom.2005.01.027>
- Breitung, J., and Eickmeier, S. (2006). Dynamic factor models. *Allgemeines Statistisches Archiv* 90(1): 27–42. <http://doi.org/10.1007/s10182-006-0219-z>
- Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: The Guilford Press.
- Clements, M.P., and Hendry, D.F. (Eds.) (2011). *The Oxford handbook of economic forecasting*. Oxford University Press.
- Cooper, R.L. (1972). The predictive performance of quarterly econometric models of the United States. In B. G. Hickman (Ed.), *Econometric Models of Cyclical Behavior* (pp. 813–948). UMI.
- Czerwiński, Z., and Guzik, B. (1980). *Prognozowanie ekonometryczne: podstawy teoretyczne i metody*. Warsaw: Polskie Wydawnictwo Ekonomiczne.
- Frale, C., Marcellino, M., Mazzi, G.L., and Proietti, T. (2010). Survey data as coincident or leading indicators. *Journal of Forecasting* 29(December 2009): 109–131.
<http://doi.org/10.1002/for.1142>
- George, E.I., and McCulloch, R.E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association* 88(423): 881–889.
<http://www.cs.berkeley.edu/~russell/classes/cs294/f05/papers/george+mcculloch-1993.pdf>
- Hansson, J., Jansson, P., and Löf, M. (2005). Business survey data: Do they help in forecasting GDP growth? *International Journal of Forecasting* 21(2): 377–389.
<http://doi.org/10.1016/j.ijforecast.2004.11.003>
- Hecq, A. (1998). Does seasonal adjustment induce common cycles? *Economics Letters* 59(3): 289–297. <http://arnop.unimaas.nl/show.cgi?fid=12365>
- Hyndman, R.J., and Koehler, A.B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22(4): 679–688.
<http://doi.org/10.1016/j.ijforecast.2006.03.001>

- Kaufmann, D., and Scheufele, R. (2013). Business tendency surveys and macroeconomic fluctuations. KOF Working Paper No. 378.
<https://www.kof.ethz.ch/en/publications/p/kof-working-papers/378/>
- Kolasa, M., Rubaszek, M., and Skrzypczyński, P. (2012). Putting the new Keynesian DSGE model to the real-time forecasting test. *Journal of Money, Credit and Banking* 44(7): 1301–1324.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4616.2012.00533.x/abstract>
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* 54(1-3): 159–178.
<http://www.sciencedirect.com/science/article/pii/030440769290104Y>
- Manski, C.F. (2014). Facing up to uncertainty in official economic statistics.
<http://www.voxeu.org/article/uncertainty-official-statistics>
- Moral-Benito, E. (2013). Model averaging in economics: An overview. *Journal of Economic Surveys, early view*. <http://doi.org/10.1111/joes.12044>
- Mycielski, J. (2010). *Ekonometria*. Warsaw: Uniwersytet Warszawski, Wydział Nauk Ekonomicznych.
- O'Hara, R.B., and Sillanpää, M.J. (2009). A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis* 4(1): 85–117.
<http://doi.org/10.1214/09-BA403>
- Próchniak, M., and Witkowski, B. (2013). Real β convergence of transition countries – robust approach. *Eastern European Economics* 51(3): 6–26.
<http://mesharpe.metapress.com/app/home/contribution.asp?referrer=parent&backto=issue,2,6;journal,1,75;linkingpublicationresults,1:106044,1>
- Reijer, A.H.J. (2012). Forecasting Dutch GDP and inflation using alternative factor model specifications based on large and small datasets. *Empirical Economics* 44(2): 435–453. <http://doi.org/10.1007/s00181-012-0560-x>
- Rubaszek, M., and Skrzypczyński, P. (2008). On the forecasting performance of a small-scale DSGE model. *International Journal of Forecasting* 24(3): 498–512.
<http://doi.org/10.1016/j.ijforecast.2008.05.002>
- Sala-i-Martin, X., Doppelhofer, G., and Miller, R.I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American Economic Review* 94(4): 813–835. <http://doi.org/10.1257/0002828042002570>
- Sims, C.A. (1980). Macroeconomics and reality. *Econometrica* 48(1): 1–48.
<http://www.jstor.org/stable/1912017>

- Stock, J.H., and Watson, M.W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics* 20(2): 147–162.
<http://doi.org/10.1198/073500102317351921>
- Ulaşan, B. (2012). Cross-country growth empirics and model uncertainty: An overview. *Economics-the Open Access Open-Assessment E-Journal* 6(2012-16): 1–69.
<http://doi.org/> <http://dx.doi.org/10.5018/economics-ejournal.ja.2012-16>
- Welfe, A. (Ed.) (2013). *Analiza kointegracyjna w makromodelowaniu*. Warszawa: Polskie Wydawnictwo Ekonomiczne.

Appendix: Description of variables used in the analysis

spingd – households final consumption expenditure index,
nakinw – investment outlays index,
eksptiu – exports of goods and services index,
imptiu – imports of goods and services index,
wdb_xxx – gross value added in xxx sector index (xxx = industry, construction, trade, transport and storage, accommodation and catering, information and communication, financial and insurance activities, real estate activities, professional and scientific activities, administrative and support service activities, public administration and defence, education, human health and social work activities, arts and entertainment, other service activities),
gus_xxx – balance of responses to question 'xxx' from a consumer sentiment survey CSO (Table A2),
gus_wb – current consumer confidence indicator (CSO),
gus_ww – leading consumer confidence indicator (CSO)
ips_wok – consumer sentiment indicator (IPSOS),
ips_kg – economic climate indicator (IPSOS),
ips_sz – advantage to make purchases indicator (IPSOS),
ips_wk – current consumer confidence indicator (IPSOS),
ips_wo – leading consumer confidence indicator (IPSOS),
zew_ies – ZEW indicator of economic sentiment,
ifo_bs – Ifo business situation indicator,
ifo_be – Ifo business expectations indicator,
biec_wwk – BIEC leading index,
biec_wpi – BIEC future inflation index,
biec_wrp – BIEC future unemployment rate index,
biec_wd – BIEC well-being index,
pmi - Purchasing Managers' index (PMI) for Polish industry,
ind_xxx - balance of responses to question 'xxx' from a business sentiment survey in industry RIED (Table A1),
hhs_xxx - balance of responses to question 'xxx' from a consumer sentiment survey RIED (Table A2),
trade - business sentiment indicator RIED in trade,
agri - business sentiment indicator RIED in agriculture,
constr - business sentiment indicator RIED in construction.

Table A1: Questions from the business sentiment survey in industry

| Symbol | Question (ind_xxs – current state, ind_xxf – projection) |
|--|--|
| ind_q1 | production |
| ind_q2 | total orders |
| ind_q3 | export orders |
| ind_q4 | stock of finished products |
| ind_q5 | prices of goods produced by enterprise |
| ind_q6 | employment |
| ind_q7 | financial standing |
| ind_q8 | Poland's macroeconomic performance |
| <i>Business sentiment survey in industry</i> , Research Institute for Economic Development, Warsaw School of Economics | |

Table A2: Questions from the consumer sentiment survey CSO & RIED

| Symbol | Question |
|---|--|
| hhs_q1, gus1 | assessment of household financial status, compared with the situation 12 months earlier |
| hhs_q2, gus2 | projected household financial status in the next 12 months |
| hhs_q3, gus3 | performance of the Polish economy in the last 12 months |
| hhs_q4, gus4 | projected performance of the Polish economy in the next 12 months |
| hhs_q5 | comparison of maintenance costs now and 12 months earlier |
| hhs_q6 | projection for the inflation rate in the next 12 months |
| hhs_q7, gus7 | projection for the unemployment rate in the next 12 months |
| hhs_q8, gus8 | an advantage to make major purchases at the present time |
| hhs_q9 | projected spending on durable consumer goods over the next 12 months in relation to the level reported in the last 12 months |
| hhs_q10 | assessment of savings and the climate for saving in the context of the country's macroeconomic performance |
| hhs_q11, gus11 | projected household's saving in the next 12 months |
| hhs_q12 | financial position of the household |
| <i>Survey of households</i> , Central Statistical Office, Research Institute for Economic Development, Warsaw School of Economics | |

Table A3: Regressors in Bayesian averaging models

| Symbol | Description |
|----------|---|
| ifo_be | Ifo business expectations indicator |
| gus2 | projected household financial status in the next 12 months |
| gus4 | projected performance of the Polish economy in the next 12 months |
| gus7 | projection for the unemployment rate in the next 12 months |
| gus11 | projected household's saving in the next 12 months |
| ips_wo | leading consumer confidence indicator (IPSOS) |
| biec_wwk | BIEC leading index |
| biec_wrp | BIEC future unemployment rate index |
| biec_wd | BIEC well-being index |
| ind_q1f | industry, production, projection |
| ind_q2f | industry, total orders, projection |
| ind_q3f | industry, export orders, projection |
| ind_q4f | industry, stock of finished products, projection |
| ind_q5f | prices of goods produced by enterprise, projection |
| ind_q6f | industry, employment, projection |
| ind_q8f | Poland's macroeconomic performance, projection |
| hhs_q1 | assessment of household financial status, compared with the situation 12 months earlier |
| hhs_q2 | projected household financial status in the next 12 months |
| hhs_q3 | performance of the Polish economy in the last 12 months |
| hhs_q4 | projected performance of the Polish economy in the next 12 months |
| hhs_q6 | projection for the inflation rate in the next 12 months |
| hhs_q7 | projection for the unemployment rate in the next 12 months |
| hhs_q9 | projected spending on durable consumer goods over the next 12 months |
| hhs_q11 | projected household's saving in the next 12 months |

Source: own estimates.

Table A4: Variables in the frequentist approach models

| Regressor | Frequentist approach without collinearity correction | | | | | | | | | | | | | | | Frequentist approach with collinearity correction | | | | | | | | | | | | | | |
|-----------|--|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| | k=0 | | | k=1 | | | k=2 | | | k=3 | | | k=4 | | | k=0 | | | k=1 | | | k=2 | | | k=3 | | | k=4 | | |
| | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I | GD P | UN E | CP I |
| ifo_be | | | 1 | | | | 1 | | | | | 1 | | | 1 | | | 1 | | | | | | | | | | | | |
| gus2 | | | 1 | 1 | | | | 1 | 1 | 1 | | | | | | | | | | | | | | | | | | | | |
| gus4 | | | 1 | 1 | | 1 | | 1 | | 1 | | | | | | | | 1 | 1 | | 1 | | 1 | | | | | | | |
| gus7 | | | | | | | | 1 | 1 | 1 | 1 | | | | 1 | | | | | | | | | | | | | | | |
| gus11 | | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | | 1 | 1 | | | | 1 | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| gus_ww | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ips_wo | | | | | | | | | 1 | 1 | | 1 | 1 | 1 | | | | | | | | | | | 1 | 1 | | 1 | 1 | 1 |
| biec_wwk | | | | 1 | | | 1 | 1 | | 1 | 1 | 1 | 1 | | 1 | | | 1 | | | 1 | | | 1 | | | 1 | | | |
| biec_wpi | | 1 | 1 | 1 | | 1 | 1 | | 1 | 1 | 1 | | 1 | 1 | 1 | | 1 | | 1 | 1 | 1 | 1 | | | 1 | 1 | | 1 | 1 | 1 |
| biec_wrp | | 1 | 1 | 1 | 1 | 1 | | | | | | 1 | | | 1 | 1 | | 1 | | 1 | 1 | 1 | | | | | | | | |
| biec_wd | | | | | | | | 1 | | | | | | | | | | | | | | | 1 | | | | | | | |
| ind_q1f | | | | | 1 | | | 1 | 1 | | 1 | | | | | | | | 1 | | | | 1 | 1 | | 1 | | | | |
| ind_q2f | | 1 | | | | 1 | | 1 | 1 | | 1 | 1 | 1 | | 1 | | | 1 | | | 1 | | | | 1 | | 1 | 1 | 1 | 1 |
| ind_q3f | 1 | | | | | 1 | | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | | | | | 1 | | | | | | 1 | | 1 | | |
| ind_q4f | | 1 | | | | | | | | | 1 | | 1 | 1 | | | 1 | | | | | | | | | 1 | | 1 | 1 | |
| ind_q5f | | 1 | | | | 1 | | 1 | 1 | 1 | | | 1 | 1 | | | | | | 1 | | | 1 | 1 | 1 | 1 | | 1 | 1 | |
| ind_q6f | | | | | | 1 | | 1 | 1 | 1 | | 1 | 1 | 1 | 1 | | | | | 1 | | | 1 | 1 | 1 | 1 | | 1 | | 1 |
| ind_q7f | | 1 | | | 1 | | | | | | | | | 1 | | | | | | 1 | | | | | | | | | | |
| ind_q8f | 1 | | 1 | 1 | | | 1 | | | | 1 | | | 1 | 1 | 1 | | 1 | 1 | | | 1 | | 1 | | 1 | | | 1 | 1 |
| hhs_q1 | | | 1 | | | | | | | | 1 | | | | 1 | | | 1 | | | | | | | 1 | | | | | |
| hhs_q2 | 1 | 1 | | 1 | | 1 | | 1 | | | 1 | | | 1 | 1 | 1 | 1 | | 1 | | 1 | | | | | | | | | |
| hhs_q4 | | | | 1 | | 1 | | 1 | | | 1 | 1 | | 1 | 1 | | | | | | | | | | | | 1 | | 1 | |
| hhs_q6 | | | | | | 1 | | | | 1 | | | 1 | | | | | | | 1 | | | 1 | 1 | | | 1 | | | |
| hhs_q7 | | | | | 1 | 1 | | | 1 | | 1 | | 1 | | 1 | | | | 1 | 1 | | | 1 | 1 | | 1 | | 1 | | |
| hhs_q9 | 1 | | | | | 1 | 1 | 1 | | | | 1 | | 1 | 1 | 1 | | | | 1 | 1 | 1 | | | | 1 | | 1 | 1 | 1 |
| hhs_q11 | | | | | | | | 1 | | | | | | 1 | 1 | | | | | | | | 1 | | | | | 1 | 1 | 1 |

Source: own estimates. “1” means that the given variable was included in the equation that describes the variability of the given indicator.

Table A5: Indicators of factors in the dynamic factor model

| Factor | Symbol | Description |
|----------|--|---|
| Factor 1 | gus1 | assessment of household financial status, compared with the situation 12 months earlier |
| | gus2 | projected household financial status in the next 12 months |
| | gus3 | performance of the Polish economy in the last 12 months |
| | gus4 | projected performance of the Polish economy in the next 12 months |
| | gus7 | projection for the unemployment rate in the next 12 months |
| | gus8 | an advantage to make major purchases at the present time |
| | gus11 | projected household's saving in the next 12 months |
| | gus_wb | current consumer confidence indicator (CSO) |
| | gus_ww | leading consumer confidence indicator (CSO) |
| | ips_wo | consumer sentiment indicator (IPSOS) |
| | ips_kg | economic climate indicator (IPSOS) |
| | ips_sz | advantage to make purchases indicator (IPSOS) |
| | ips_wk | advantage to make purchases indicator (IPSOS) |
| | ips_wo | leading consumer confidence indicator (IPSOS) |
| | biec_wrp | BIEC future unemployment rate index |
| | biec_wd | BIEC well-being index |
| | ind_q5f | prices of goods produced by enterprise |
| | hhs_q1 | see gus1 |
| | hhs_q2 | see gus2 |
| | hhs_q3 | see gus3 |
| | hhs_q4 | see gus4 |
| | hhs_q7 | see gus7 |
| hhs_q8 | see gus8 | |
| hhs_q9 | projected spending on durable consumer goods over the next 12 months | |
| hhs_q10 | assessment of savings and the climate for saving | |
| hhs_q11 | see gus11 | |
| Factor 2 | pmi | Purchasing Managers' Index (PMI) for Polish industry |
| | ifo_bs | Ifo business situation indicator |
| | ifo_be | Ifo business expectations indicator |
| | ind_q1s | industry, production, current state |
| | ind_q1f | industry, production, projection |
| | ind_q2s | industry, total orders, current state |
| | ind_q2f | industry, total orders, projection |
| | ind_q3s | industry, export orders, current state |
| | ind_q3f | industry, export orders, projection |
| | ind_q6s | industry, employment, current state |
| | ind_q6f | industry, employment, projection |
| | ind_q7s | industry, financial standing, current state |
| | ind_q7f | industry, financial standing, projection |
| | ind_q8s | Poland's macroeconomic performance |
| | ind_q8f | Poland's macroeconomic performance |
| | constr | business sentiment indicator RIED in construction |

Table A5 continued

| | | |
|----------|--|--|
| Factor 3 | zew_ies ifo_bs gus1 gus2 biee_wwk biee_wpi biee_wrp ind_q1f ind_q2f ind_q3f ind_q4s ind_q4f ind_q5f hhs_q9 hhs_q12 | ZEW indicator of economic sentiment Ifo business situation indicator Assessment of household financial status, compared with the situation 12 months earlier projected household financial status in the next 12 months BIEC leading index BIEC future inflation index BIEC future unemployment rate index industry, production, projection industry, total orders, projection industry, export orders, projection industry, stock of finished products, current rate industry, stock of finished products, projection prices of goods produced by enterprise, projection projected spending on durable consumer goods over the next 12 months financial position of the household |
|----------|--|--|

Source: own estimates.

Figure A1

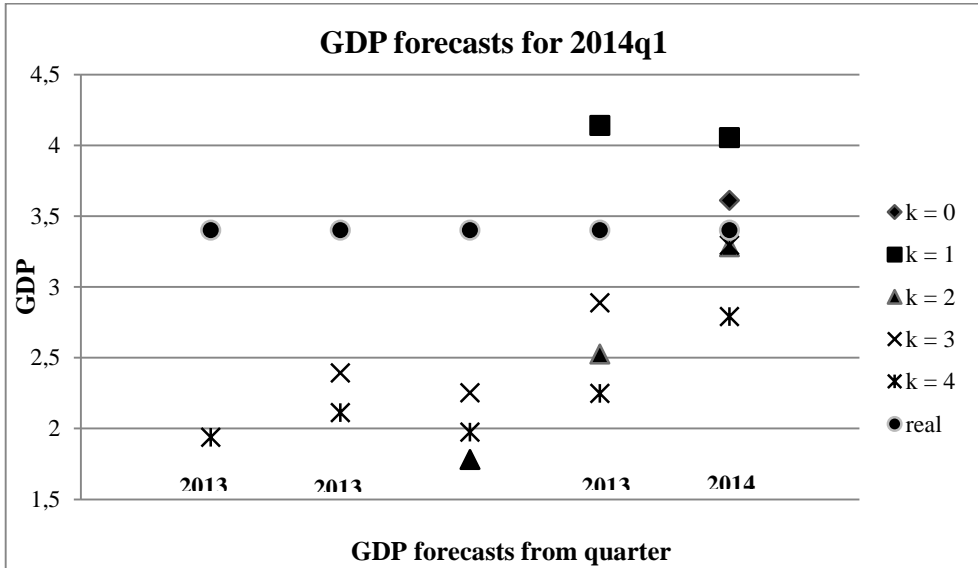


Figure A2

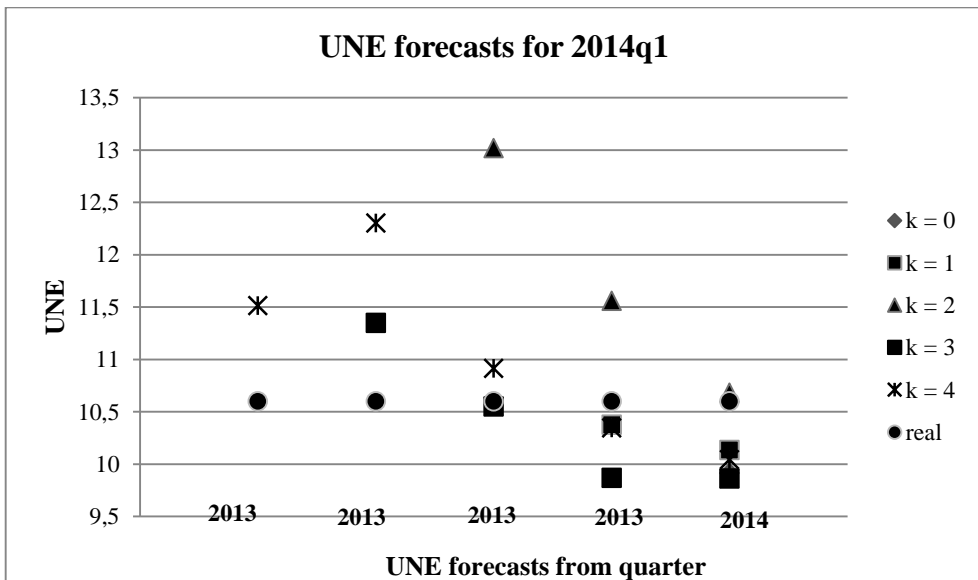
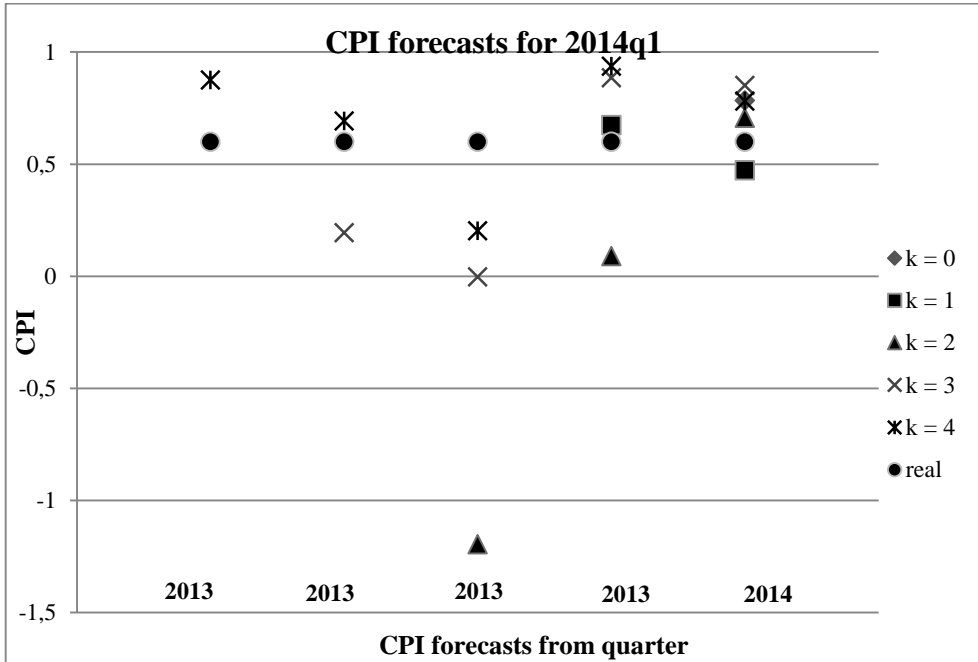


Figure A3



Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

<http://dx.doi.org/10.5018/economics-ejournal.ja.2015-31>

The Editor