LIETUVOS BANKAS

WORKING PAPER SERIES

No 13 / 2012

GENERATING SHORT-TERM FORECASTS OF THE LITHUANIAN GDP USING FACTOR MODELS

by Julius Stakėnas

WORKING PAPER SERIES
No 13 / 2012

GENERATING SHORT-TERM FORECASTS OF THE LITHUANIAN GDP USING FACTOR MODELS

by Julius Stakėnas*

* Bank of Lithuania. The author is grateful for two anonymous referees for their helpful comments. For correspondence: Economics Department, Bank of Lithuania, Totorių g. 4, LT-01121 Vilnius, Lithuania; e-mail: jstakenas@lb.lt.

Address
Totorių g. 4
LT-01121 Vilnius
Lithuania
Telephone (8 5) 268 0125
Fax +370 5 212 4423

Internet
http://www.lb.lt

Working Papers describe research in progress by the author(s) and are published to stimulate discussion and critical comments.

The Series is managed by Economic Research Division of Economics Department

ISSN 2029-0446 (ONLINE)

Contents

Abstract

This paper focuses on short-term Lithuanian GDP forecasting using a large monthly dataset. The forecasting accuracy of various factor model specifications is assessed using the out-of-sample forecasting exercise. It is argued that factor extraction by using a simple principal components method might lead to a loss of important information related GDP forecasting, therefore, other methods should be also considered. Performance of several factor models, which relate the factor extraction step to GDP forecasting, was tested. The effect of using weighted principal components model, with weights depending on variables' absolute correlation with GDP, was explored in greater detail. Although factor models performed better than naive benchmark forecast for GDP nowcasting and 1 quarter ahead forecasting, we were unable to set up the ranking among different factor model specifications. We also find that a small scale factor model with 5 variables (which could be regarded as the most important monthly variables for GDP nowcasting) is able to nowcast GDP better than models with a full data set of 52 variables, which might indicate that for the case of the Lithuanian economy, a smaller scale factor models may be more suitable.

Keywords: GDP forecasting, factor models, principal components
JEL classification: C22, E37

Santrauka

Straipsnyje lyginamas Lietuvos BVP prognozių tikslumas naudojant įvairių specifikacijų faktorinius modelius didelės apimties mėnesinių duomenų rinkiniams. Modeliai vertinami atliekant pseudo realaus laiko prognozavimo pratimą ir lyginant prognozes su jų realizacijomis. Straipsnyje keliama mintis, kad konstruojant faktorius pagrindinių komponenčių metodu, galimai prarandama BVP prognozavimui svarbi informacija. Siekiant ištaisyti šį trūkumą, į prognozavimo pratimą įtraukiamos kelios faktorinių modelių specifikacijos, susiejančios faktorių konstravimą su BVP prognozavimu. Plačiau aptariamas svertinis pagrindinių komponenčių modelis, suteikiantis kintamiesiems skirtingą svarbą formuojant faktorius priklausomai nuo kintamojo absoliutinės koreliacijos su BVP augimu. Gauti rezultatai leidžia teigti, kad faktoriniai modeliai, einamojo ir kito ketvirčio Lietuvos BVP augimą prognozuoja tiksliau nei kontrolinis modelis, tačiau negalime vienareikšmiškai suranguoti faktorinių modelių specifikacijų pagal jų prognozių tikslumą. Pažymėtina, kad faktorinis modelis, naudojantis tik 5 ekonomiškai svarbius kintamuosius, prognozavo geriau nei modeliai naudojantys 52 kintamuosius, ir tai rodo, kad Lietuvos duomenims mažesnės apimties faktoriniai modeliai, tikėtina, yra labiau taikytini.

Raktiniai žodžiai: BVP prognozavimas, faktoriniai modeliai, pagrindinės komponentės
JEL classification: C22, E37

1. Introduction

The abundance of information, which is nowadays available to an econometrician, may not only benefit with more accurate forecasts of economic variables, but also poses a challenge of how to effectively use it. Standard modelling techniques, such as regression and VAR, may not be applicable for modelling a large number of variables, because there would be too many parameters to estimate. With a growing number of variables, included into equations, these estimates quickly become unreliable or even unfeasible. Some kind of dimension reduction technique is needed and this is where factor models may come very handy.

The starting point of the factor analysis is the observation that there is usually only a small number of common factors responsible for a big part of variation in many economic variables. If we could extract these factors from the observed data, they could be useful for forecasting, as they would summarize information from many different sources.

Various factor models have been proposed in the literature to exploit large datasets for forecasting purposes. Stock and Watson (2002) used principal components to extract factors from the data. Forni, Hallin, Lippi and Reichlin (hereafter FHLR) (2003) developed the so-called generalized dynamic factor model, estimating common and idiosyncratic covariance matrices through dynamic principal components and using these estimates to get generalized principal components. Factor dynamics were incorporated by Giannone, Reichlin, Sala (2004) putting principal components into state space modelling framework.

The aforementioned factor models' variants were tested on numerous cases, using datasets collected from many different countries. A fine example is the study of Barhoumi et al. (2008), which examines performance of the factor model based GDP forecasts for 9 European countries and the euro area as a whole. Although in this study factor model forecasts outperformed benchmark forecasts for the euro area and for six euro area countries separately, it failed to produce accurate forecasts for 3 new Member States (including Lithuania). This partly was a motivation to take a second look at Lithuania's case, operating longer data series.

Another reason to review previous findings on the factor model based GDP forecasts for the Lithuanian data was the intention to examine ways to relate the factors' extraction step to their later use for GDP forecasting. This aspect was neglected in the study of Barhoumi et al. To achieve this goal, we adopt several techniques, producing so called supervised factor models (Lee, Tu (2009)). Following Bai and Ng (2008), we use LARS algorithm to select a set of informative variables from a larger set. The selected set is later used to obtain factors. Another method examined is partial least squares factor model, introduced by Wold (1996). This method differs from unsupervised factor models, as it uses a different objective function in the factor extraction step. Lastly, we examine different weighting schemes, assigning importance weights for variables and performing weighted principal components analysis. Usage of weighted principal components, with weights being a function on variables' absolute correlations with GDP, is not common in economic literature, however we feel that favoring some variables over the others is quite an intuitive approach. Some potential benefits and drawbacks arising from such weighting are also examined.

The paper proceeds as follows. In section 2 we introduce factor models which we will use in our forecast evaluation exercise. Section 3 discusses ways to "supervise" factor extraction to obtain factors which would be more useful for GDP forecasting. Section 4 presents the design of our forecast evaluation exercise. The main empirical findings are presented and discussed in section 5. Finally, section 6 concludes.

2. Factor models and their estimation

This section reviews the formulation and estimation of three popular factor models used in the literature for forecasting purposes: principal components, generalized principal components and state space model.

A factor model, used for forecasting is a system of two equations:

$$\begin{cases} X_t = \chi_t + \xi_t = \Lambda f_t + \xi_t \\ y_{\tilde{t}}^Q = \mu + \beta' f_{\tilde{t}}^Q + \varepsilon_{\tilde{t}}^Q \end{cases} \tag{1}$$

where:

$X_t = (x_{1t}, x_{2t} \dots x_{nt})'$ $- n \times 1$ variable vector,

$f_t = (f_{1t}, f_{2t} \dots f_{mt})'$ $- m \times 1$ factor vector, $m < n$,

$\xi_t = (\xi_{1t}, \xi_{2t} \dots \xi_{nt})'$ $- n \times 1$ idiosyncratic component vector,

$\chi_t - n \times 1$ common component vector,

$y_{\tilde{t}}^Q$ − quarterly GDP variable,

$f_{\tilde{t}}^Q$ $- m \times 1$ quarterly factor vector,

$\Lambda -$ $n \times m$ loading matrix,

$t = 1 \dots T$,

$\tilde{t} = 1 \dots \tilde{T}$,

$E(\xi_{it}\chi_{ls}) = 0, \ \forall \, i, l = 1 \dots T, \ t = 1 \dots n, \ s = 1 \dots m$,

ξ_{lt} and ξ_{kt} can be correlated to "some extent".

The first equation in the system (1) states that variable vector X_t can be decomposed into common component χ_t and idiosyncratic component ξ_t. X_t is explained by factors f_t, which are orthogonal to ξ_t and are driven by variable specific shocks. In our model, the first equation consists of monthly variables, whereas factors in the second equation are of the quarterly frequency, to relate quarterly GDP and factors through simple OLS. Hence, the modelling scheme is fairly simple: firstly we extract monthly factors from variable vector X_t, then we form quarterly factors (we will elaborate on the aggregation of factors in section 4.2) and use them as explanatory variables in simple regression to model GDP variation. The subsequent factor models, reviewed in the paper, differ only in the factor extraction step.

2.1 Principal components

Principal components (PC) estimator aims at maximizing variances of factors f_{it}, while constraining them to be orthogonal, to make sure newly received information is different from already available. In other words, the first extracted factor has the highest variance possible and each subsequent factor has its variance maximized subject to a constraint that its projection is orthogonal to all previous projections. The constraint for the projection to be of norm equal to 1, restricts the objective function to always remain finite. The procedure can be expressed as follows:

$$\begin{cases} f_j = Xw_j, \quad j = 1 \dots m \\ \max\limits_{w_j} \left(f_j' f_j\right) = \max\limits_{w_j} (w_j' X' X w_j) \\ w_j' w_j = 1 \\ w_j' w_l = 0, \; l = 1 \dots j-1 \end{cases} \tag{2}$$

where:

$X = (X_1, X_2 \dots X_T)'$,

$w_j - n \times 1$ vector of weights used for factor j.

This constrained optimization can be performed using Lagrange method, which in turn leads to computing m eigenvectors[1] of $X'X$ matrix, corresponding to m largest eigenvalues.

By performing the maximization of a factor's variation, we try to retain as much information contained in X as possible, while at the same time reducing X dimension. The standardization of variables in X must be performed before the maximization in order to avoid scale and mean effect on the computation of factor loadings (variables with higher mean or variance would be appointed with bigger weights forming a factor). This sensitivity to scale is employed in the generalized principal components and weighted principal components methods, which may lead to better forecasts of GDP growth.

An alternative (and probably more intuitive) way to find principal components is through minimizing the sum of squared idiosyncratic components:

$$\begin{cases} \min\limits_{\Lambda, f_1, f_2 \dots f_T} \sum_{t=1}^{T} [(X_t - \Lambda f_t)'(X_t - \Lambda f_t)] \\ \Lambda' \Lambda = I_m \end{cases} \tag{3}$$

where $\Lambda = (w_1, w_2 \dots w_m)$.

It can be shown that maximization and minimization problems are equivalent.

2.2 Generalized principal components

Generalized principal components (GPC) method "generalizes" PC in an analogous manner as GLS "generalizes" OLS. If we had information that i-th variable has high idiosyncratic dispersion $\sigma_{\xi i}^2$ and is not very reliable for factor estimation, we would like the variable to receive smaller weights w_{ij}.

FHLR (2005) implement this idea by choosing variable weights which maximize the common-to-idiosyncratic variance ratio. In other words, we keep the dispersion of an idiosyncratic "factor" equal to 1 and maximize the dispersion of common factor.

Such constrained optimization is defined by the system[2]:

[1] While there exist criteria in the literature to determine the number of factors (see e.g. Bai and Ng (2002)), in our study, models with 2 factors yield the best results. Therefore, here and hereafter we assume $m = 2$.

$$\begin{cases} f_j = Xw_j, \quad j = 1 \dots m \\ \max\limits_{w_j}(f_j' f_j) \;=\; \max\limits_{w_j}(w_j' X' X w_j) \\ w_j' \Sigma_\xi w_j = 1 \\ w_j' \Sigma_\xi w_l = 0, \;\; l = 1 \dots j-1 \end{cases} \tag{4}$$

where Σ_ξ is a covariance matrix of idiosyncratic component.

In general, we do not know the value of matrix Σ_ξ, therefore, GPC estimator is infeasible. To estimate Σ_ξ, we use the method proposed by FHLR (2003), employing the estimation of dynamic principal components. As in FHLR (2003), we use only diagonal entries of the estimate $\hat{\Sigma}_\xi$, setting non-diagonal elements equal to 0.

It can be shown that the problem (4) leads to finding m eigenvectors of matrix $(\Sigma_\xi^{-1/2} X' X \Sigma_\xi^{-1/2})$, corresponding to m largest eigenvalues. As we used only diagonal entries of Σ_ξ, it is easy to interpret elements of Σ_ξ as reflecting our view of variable significance when forming a factor (variables corresponding to bigger Σ_ξ elements are more likely to get small weights w_{ij}). We will later use this interpretation for weighted principal components method to distinguish variables, which in our view are more significant for GDP forecasting than others.

2.3 State space model

The two principal component models described previously are actually dynamic factor models, as extracted factors can generally be lags of current factors. However, dynamics of factors was not parameterized or estimated in the previous models. To model factors' relationships, a state space factor model expands principal components model with a VAR equation for factors:

$$\begin{cases} X_t = \Lambda F_t + \xi_t \\ F_t = AF_{t-1} + \eta_t \end{cases} \tag{5}$$

where:

$Cov(\xi_t, \xi_t) = \Sigma_\xi$,

$\eta_t \sim N(0, \Sigma_\eta)$,

$Cov(\xi_t, \eta_s) = 0, \quad \forall\, t, \forall s$,

Σ_η, Σ_ξ – diagonal covariance matrices.

Model (5) is estimated using a two-step estimation method described in Doz, Giannone, Reichlin (2006). To understand the estimation of (5), we firstly concentrate on model (6), which has equivalent assumptions as (5):

$$\begin{cases} X_t = \hat{\Lambda} f_t + \xi_t \\ f_t = A_1 f_{t-1} + A_2 f_{t-2} + \dots + A_p f_{t-p} + \eta_t \end{cases} \tag{6}$$

[2] We use total covariance matrix, instead of common covariance matrix, for comparability with simple PC, as it does not influence results.

In the first step, the first equation of (6) is estimated by principal components, producing estimates of $\widetilde{\Lambda}$ and f_t. Factors f_t are then subsequently used to estimate the second equation of (6) by OLS. Now we rewrite the estimated model (6) in the form of (5). Once the parameters of model (5) are estimated, in the second step factor values are filtered and smoothed to obtain new estimates of factors F_t. As it is pointed out in Doz, Giannone, Reichlin (2006), the procedure of reestimating A, using fresh estimates of F_t and then smoothing again to get new F_t can be repeated many times producing more efficient estimates. In our case, we performed Kalman smoothing twice.

3. Factor model refinements

Although factor models are praised for their ability to employ large datasets for modeling purposes, it is not obvious whether using more information will always yield better estimates or forecasts. As it is shown in Boivin and Ng (2006), using more data and constructing factors from variables with highly correlated idiosyncratic components can result in less efficient factor estimates. Also, if forecasting power is provided by a factor which is dominant in a small sample, but dominated in a larger sample, obviously, using more data would lead to inferior results.

As we have seen in section 2, the standard factor extraction procedure does not take into account factors' later use for GDP forecasting. The main objective is to reduce data dimensionality, retaining as much of its information as possible. Thus, the factors will only reflect the data, "averaging" them in a specific way and likely losing some valuable information about the forecast variable.

In this section we will examine 3 methods (variable selection, partial least squares and weighted principal components), relating factors' extraction with their purpose of GDP forecasting. We will refer to these methods as supervised factor models.

3.1 Variable selection

The idea of the variable selection is simple: instead of using the whole set of variables for factor estimation, we may select a smaller set, consisting only of those variables which are in some sense most relevant for forecasting. In this way, we omit variables which could potentially hinder forming factors with good forecasting properties. The selected variables are often called "targeted predictors" (see Bai and Ng (2008) for a more detailed study).

We employed the least angle regression with elastic net (LARS–EN) to perform the variable selection. This method uses LARS algorithm described in Efron et al. (2004) satisfying EN criteria (see Zou and Hastie (2005)).

LARS algorithm starts with an empty set, selecting a variable having the highest absolute correlation with GDP. In each of the subsequent steps, we modify the GDP variable, subtracting part of information contained in the selected set and selecting a new variable, having the highest absolute correlation with the modified GDP. In the process of a new variable selection, LARS algorithm takes into account that some of the information the new variable contains is already reflected by other set variables. At the same time, the algorithm is still able to select mutually correlated variables, unlike, for example, forward selection method.

The results reported in section 5 were produced by selecting 20 variables and using penalty parameter value equal to 0.5. Performance of the variable selection with other parameter values (1, 1.5) and sizes of a variable set (10, 40) were also tested, however, the results did not change substantially, therefore, for brevity reasons, in section 5 we present results only for one pair of parameters.

3.2 Partial least squares

To relate GDP forecasting and factor extraction step using the partial least squares (PLS) model, we aim to extract orthonormal factors, having the highest square covariances with GDP (y^Q). Hence, the objective function looks as follows:

$$\max_{w_i} cov\left(f_i^Q, y^Q\right)^2$$

where:[3]

$i \quad = 1 \ldots m,$
$X_t^Q = \Lambda f_t^Q + \xi_t^Q,$
$\Lambda \quad = (w_1, w_2 \ldots w_m),$
$f_t^Q \;= (f_{1,t}^Q, f_{2,t}^Q \cdots f_{m,t}^Q)',$
$\{f_i^Q, \;\; i = 1 \ldots m\}$ – set of orthonormal projections.

The first projection f_1^Q and weights w_1 are found by maximizing:

$$\max_{|w_1|=1} cov(X^Q w_1, y^Q)^2 = \max_{|w_1|=1} \left(w_1' X^{Q'} y^Q y^{Q'} X^Q w_1\right)$$

This maximization is equivalent to finding unit eigenvector of matrix $X^{Q'} y^Q y^{Q'} X^Q$, corresponding to largest eigenvalue.

To obtain subsequent factors/projections, orthogonal to the previously found, we deflate data matrix X^Q by:

$$X^Q := X^Q - f_1^Q w_1'$$

The maximization procedure is then repeated. Again, we find unit eigenvector of $X^{Q'} y^Q y^{Q'} X^Q$, corresponding to largest eigenvalue. We repeatedly deflate X^Q and maximize objective function to extract all m needed factors.

Contrary to PC model, PLS model relates factors' construction to their use for forecasting GDP. As it is pointed out in Rosipal and Krämer (2006), PLS can be perceived as an intermediate model between principal components and canonical correlation analysis (CCA). The reasons for such interpretation are clearer after the examination of respective objective functions:

$$CCA\colon \max_{|w|=1} corr(X^Q w, y^Q)^2$$
$$PC\colon \max_{|w|=1} var(X^Q w)$$
$$PLS\colon \max_{|w|=1} cov(X^Q w, y^Q)^2 = \max_{|w|=1} var(X^Q w) corr(X^Q w, y^Q)^2 var(y^Q)$$

[3] We use superscript Q for X and factors f to denote we are dealing with quarterly variables.

Hence, with PLS we try to reach a compromise between goals of PC and CCA, maximizing the product of their objective functions.

3.3 Weighted principal components

It is obvious that some variables are more important for GDP forecasting than others: e.g the Lithuanian quarterly retail sales variable has a correlation coefficient with GDP equal to 0.72, while the producer price index's correlation with GDP is only 0.08. Also, it should be noted that usefulness of a specific variable changes over different forecasting horizons. For example, whilst variables of industrial output and foreign trade are followed for GDP nowcasting, various confidence indicators and survey data can indicate movements of future GDP. However, factor models treat all variables in X equally, having the objective to retain as much information contained in data matrix X as possible. Although we may use variable selection and PLS methods to try to deal with this drawback, another more straightforward alternative would be to assign certain weights to variables which would reflect our view of variable usefulness regarding GDP forecasting.

This approach leads to the weighted principal components method (WPC). In economic literature, WPC method is mainly used to weight variables according to their idiosyncratic error variances (GPC method). In this study, we will examine the performance of WPC, when weights are formed as a function of a variable's absolute correlation with GDP. Similar interpretation of WPC can be found in the image recognition literature (see e.g. Thomaz et al. (2010)).

As it is pointed out in section 2, in advance of factor extraction for the PC model, X is standardized to avoid scale effect (otherwise, variables with larger variances would be favoured and appointed with larger weights in forming a factor). We shall use this "channel" to transfer our view of a variable's utility for forecasting. We modify variances of standardized variables in X, multiplying X by diagonal weights' matrix D:

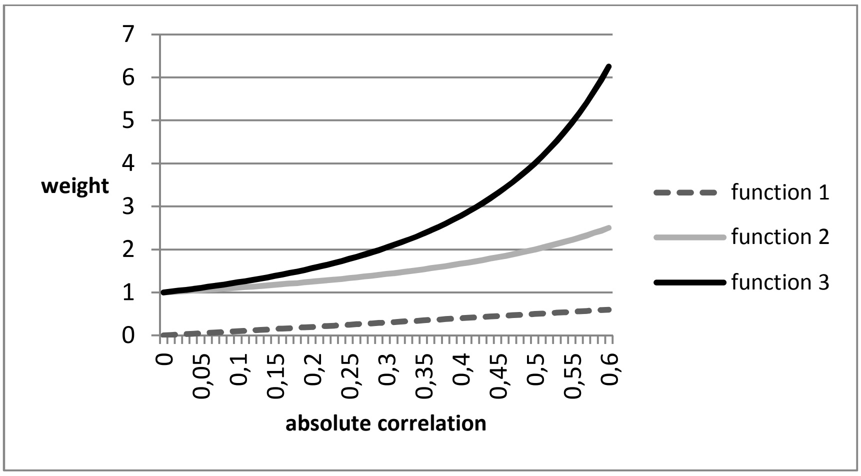
$$X := XD, \quad D = diag(d_1, d_2 \dots d_n)$$

Now we can apply factor models, discussed previously, for the transformed X.

It should be noted that when using weighted principal component analysis, factors' interpretation changes. We can no longer interpret factors as a few latent variables driving the variation of economical variables, because factors now compromise (similarly to PLS) between reflecting X's variability and information about GDP.

The values of significance weights d_i are generally unknown, therefore, this approach is prone to subjectivity. To add some objectivity to the issue, we model d_i as a function of absolute correlation with GDP, i.e. $d_i = f(|corr(X_i^Q, y^Q)|)$. Intuition also suggests that function $f(|corr(X_i^Q, y^Q)|)$ may be nonlinear. Variables having especially high absolute correlation with GDP (the number of such variables is small for the Lithuanian data) may possess some rare additional information about GDP variation and should be rewarded with higher weights than linear function would appoint. The graphs of functions, examined for variable weighting, are presented in Figure 1.

Figure 1. Variable significance weights depending on absolute correlation with GDP



The functions are defined as:

$$f_1 = \left|corr\left(X_i^Q, y^Q\right)\right|, \quad f_2 = \frac{1}{1-\left|corr(X_i^Q, y^Q)\right|}, \quad f_3 = \frac{1}{1-\left|corr\left(X_i^Q, y^Q\right)\right|^2} \tag{7}$$

We find that in our case weighting function f_3 produces smallest forecasting errors, therefore, the results for WPC in section 5 are presented using f_3. On the other hand, if one preferred forecasts, which are likely less dependent on individual variables, functions f_1 and f_2 would be a better choice.

To compare WPC with other methods, discussed previously, it is useful to have another look at their objective functions:

PC: $\max_{|w|=1}(w'X'Xw)$

GPC: $\max_{|w|=1}\left(w'\Sigma_\xi^{-1/2}X'X\Sigma_\xi^{-1/2}w\right)$

PLS: $\max_{|w|=1}(w'D\,1_{n\times n}Dw)$

WPC[4]: $\max_{|w|=1}(w'DX'XDw)$

where:

D $=$ $diag(corr\left(X_1^Q, y^Q\right), corr\left(X_2^Q, y^Q\right) \ldots corr(X_n^Q, y^Q))$

$1_{n\times n}$ – $n \times n$ matrix of ones

The construction of objective functions suggests that WPC pays more attention to X covariances than PLS, which uses the matrix of ones instead of $X'X$. Also, we can see that GPC is a WPC method, in which variables are weighted by inverse of their idiosyncratic components' standard deviations.

For comparability reasons, in our empirical study variable weighting is applied to GPC and state space models as well. In what follows, for the GPC model we use double weighting – first

[4] To make WPC and PLS comparable, function f_1 is used for WPC.

time variables are weighted according to their idiosyncratic components' errors and second time according to their correlations with GDP.

4. Forecast evaluation exercise

This section is devoted to description and details of forecast evaluation exercise, used to test performance of different factor models for the Lithuanian data.

4.1 Forecasting in pseudo real-time

In the previous sections, our main focus was on factor extraction techniques, whereas little was told about how the forecasts are actually obtained. We now turn our attention to this matter.

Once we extracted monthly factors $\{f_{i,t},\ i = 1 \ldots m,\ t = 1 \ldots T\}$, we then use them (or now known loading matrix Λ) to form quarterly factors $\{f_{i,t}^Q, i = 1 \ldots m,\ t = 1 \ldots \tilde{T}\}$. Quarterly factors are incorporated in a simple OLS regression on quarterly GDP:

$$y_t^Q = \alpha_0 + \alpha_1 f_{1,t}^Q + \alpha_2 f_{2,t}^Q + \cdots + \alpha_m f_{m,t}^Q + \varepsilon_t^Q \tag{8}$$

It is straightforward to compute forecasts y_{t+h}^Q, when values of factors f_{t+h}^Q are known, as we simply insert f_{t+h}^Q into equation (8). For the state space model, this will be always the case, as we can forecast factors using Kalman filter equations. However, we cannot forecast factors with PC or GPC model, as factor dynamics is not specified in these models. For these cases, we evaluate regressions (9) for every forecast horizon h:

$$y_{t+h}^Q = \alpha_0 + \alpha_1 f_{1,t}^Q + \alpha_2 f_{2,t}^Q + \cdots + \alpha_m f_{m,t}^Q + \varepsilon_{t+h}^Q \tag{9}$$

To obtain forecasts y_{T+h}, we insert last available factor values $f_{i,T}^Q$ into (9).

It should be noted that, for variable selection method, factors $f_{i,t}^Q$ in equation (9) are formed using a set of variables which best explain y_{t+h}^Q, hence, for every h we extract different factors. Similarly, when the WPC model is applied (the same arguments are valid for the PLS model as well), weights are formed as a function of absolute correlations $|corr(X_{i,t}^Q, y_{t+h}^Q)|$, i.e. factors $f_{i,t}^Q$ are different in equation (9) for every h. This feature helps to better exploit the structure of the data. Generally, we can group variables into leading, coincident and lagging indicators, thus, it is intuitive that for different h in the (8) equation, weights of variables comprising factors $f_{i,t}^Q$ should differ. WPC model works exactly this way, assigning different significance weights for different h, while variable selection method simply uses different sets of variables for every h. It should be also noted that when variable weighting is applied for the state space model, weights are formed using only correlations with $h = 0$, as the model is designed to extend factors to the future. However, this results in not fully exploiting the structure of the data – variable weighting always favours only coincident indicators.

Data used in the study spans the period of Q1 1996 – Q3 2011. Of this period, we use Q2 2000 – Q1 2011 for out-of-sample forecast evaluation. For every month in Q2 2000 – Q1 2011, we estimate GDP of a previous quarter[5], current quarter and two quarters ahead. Models are reestimated every month before computing forecasts. Due to differences in the publication lag structure (e.g. industrial output data is usually published later than the survey data), we use a pseudo real-time forecast design as in Barhoumi et al. (2008). 15 09 2011 is taken as a reference date and we assume that for all the previous months from our sample, publication lags on the 15-th day of a month are the same as they were in 15 09 2011. Data which should not be known at a time of a specific forecast is deleted, thus simulating real time forecasts.

The whole forecast evaluation exercise was implemented using Matlab software.

4.2 Variable transformations and aggregation of factors

In order to use variables in the factor extraction step, firstly they must be stationarized. We find that month-on-month difference transformation is suitable for most of the variables, however, it is not obvious what kind of transformations (if any) should be applied to confidence indicator variables. We believe that confidence levels and confidence changes possess different information and both might be useful for GDP forecasting. On the other hand, differences in confidence level may not be very beneficial for forecasting during the times of low variation in confidence level. Due to this drawback, we chose to work with confidence level variables. Aside from difference transformation, some variables were also seasonally adjusted and logarithmically transformed. More details on specific variable transformations can be found in Table 9 of Appendix D.

As stated previously, in our modelling procedure initially monthly factors $\{f_{i,t},\ i = 1 \dots m,$ $t = 1 \dots T\}$ are extracted and then they are aggregated to a quarterly frequency to obtain $\{f_{i,t}^{Q}, i = 1 \dots m,\ t = 1 \dots \tilde{T}\}$. However, in our case, when month-on-month variable transformations are used together with levels, we are not able to filter monthly factors to obtain quarterly factors with a desirable property that factors are linear combinations of quarterly variables. When aggregating monthly factors, we chose between two alternatives. Firstly, we may treat factors as new monthly variables not paying attention to their inner structure and aggregate them by simple averaging. Secondly, we may assume that weights computed from monthly variables are also valid for quarterly variables and obtain F_t^Q as $F_t^Q = X_t^{Q'} w$. In our study, the state space models factors were aggregated through the averaging of monthly factors, whereas for other methods, the second alternative was chosen[6] – quarterly variables X_t^Q were used and weights were computed from the monthly variable set.

4.3 Data related issues

Factor modelling is a theory-free modelling technique, as we use our economic knowledge only at the outset – when choosing a set of monthly variables to include into the model. In our view,

---

[5] Statistics Lithuania announce the first GDP estimate of a quarter on the 28th day of the following quarter, hence backcasting is performed only on the first month of a quarter.

[6] We could also have computed weights from aggregated quarterly variables X_t^Q but this would mean we would be using 3 times less data for weights' computation.

this decision may also have a significant effect on forecasts, therefore we chose our dataset quite carefully, obtaining a rather small dataset, as compared to datasets used in factor modeling literature.

The whole list of monthly variables selected for the study can be found in Appendix D and consists of 19 survey variables, 6 industry production variables, 9 trade data variables, 5 price variables, 6 financial variables and 7 variables of other type (52 indicators in total). The number of variables tested in the model is actually greater than 52, as some variables were discarded due to their noisy nature, negligible economic impact or undesirable empirical relations with GDP which are unlikely to hold in the longer term.

An important issue in factor modeling is dealing with the so-called "ragged edges" of data caused by variables' different publication lags. Also, the latest available data usually do not span the whole quarter, hindering the aggregation of variables to quarterly frequency. To be able to extract factors and use them for forecasting, we need to balance the data and to fill in the last quarter. Although we can use the whole variable set for data balancing, it seems that in our case various multivariate and factor models do not provide much additional forecasting power over univariate models. Therefore, we resort to AR(3) models for data balancing and forecasting. In the case of the state space model, we act differently, as missing X_t observations are provided by Kalman filter equations.

5. Results

In order to compare results of the forecast evaluation exercise described in section 4, we use the RMSE (root mean square error) and MAE (mean absolute error) criteria. The criteria are presented as ratios to the corresponding criteria of benchmark forecasts. Thus, the value of criterion below 1 would suggest that forecasts of a certain factor model are more accurate than the ones of the benchmark model. For the benchmark model, we took a naïve approach, when forecasts of GDP for the forthcoming quarters are computed as an average of a past period. Results of the forecasting exercise can be found in Tables 2 and 3 of Appendix A.

5.1 General results

The main results of the forecasting exercise may be summarized as follows:

- Our three main factor models (GPC, PC and state space model) produced rather similar results and we were unable to discern the best performance in terms of forecasting accuracy.

- The variable selection has not improved forecasts of the unsupervised[7] models.

- The PLS method seems to be the least suitable alternative of all tested factor models in the analysed case.

[7] Here and hereafter we refer to "unsupervised" models, as models which do not relate factor extraction with factors' purpose for forecasting.

- Models with variable weighting (WPC) slightly improved forecasts of the unsupervised models, but only for the very short term forecasts (nowcasting from the first and the second month of the running quarter).

- WPC and variable selection methods produced rather unstable 1-quarter and 2-quarter ahead forecasts.

- None of the factor models were able to predict the extent and timing of the GDP drop during the 2008-2009 crisis.

5.2 WPC and PC model comparison

We now take a closer look at the differences between the PC and WPC model forecasts and the reasons behind these differences. The graphs of the forecasts for both models can be found in Appendix B.

Looking at the forecast graphs, we can see that WPC forecasts are considerably more volatile than PC forecasts. This difference in volatility is especially visible in longer-term forecasts. Despite the differences, it seems that none of the models are suitable for longer-term forecasting – PC longer-term forecasts become reminiscent of a straight line, while WPC forecasts often miss the actual GDP values and produce large forecast errors.

For further model comparison, let us have a look at some of the factor weights assigned to variables by WPC and PC models. Forecasts are produced using the following equation:

$$y_{t+h}^{Q} = \hat{\alpha}_{0,h} + \hat{\alpha}_{1,h} f_{1,t}^{Q} + \hat{\alpha}_{2,h} f_{2,t}^{Q} \tag{10}$$

We computed PC and WPC variable weights for $h = 0, 1, 2$ in the equation (10). The results for the first 5 most important variables in a factor can be found in Tables 5 – 8 in Appendix C.[8]

We shall stress two aspects of PC and WPC differences in weights:

- WPC weights exhibit larger variation as compared to PC.
- WPC and PC weights reflect differences in their objective functions.

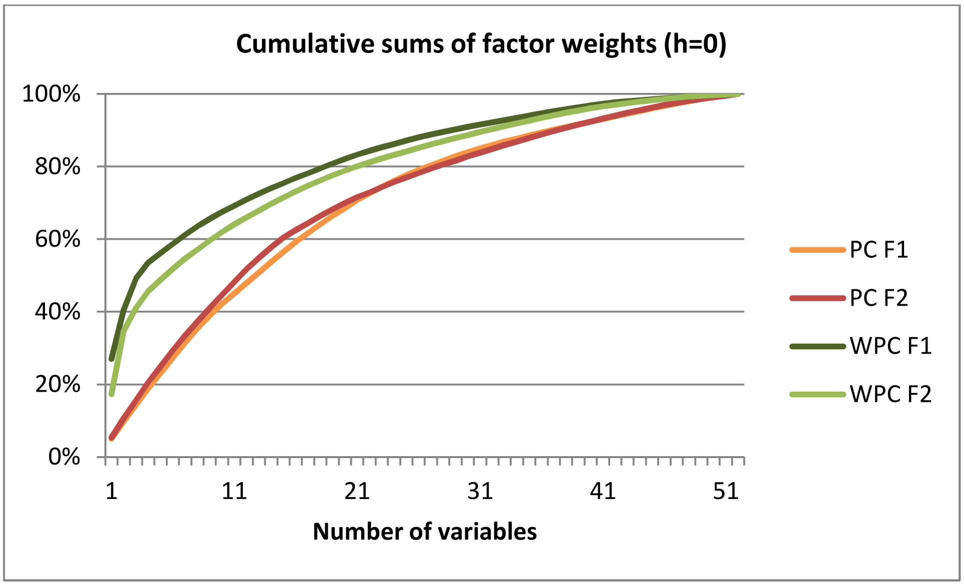
Comparing Tables 5 and 6, we can see that the WPC model produces considerably larger differences in factor weights than the PC model. Actually, only 3 variables in the WPC model (for $h = 0$), are responsible for 49% of F_1 variation, while in the case of the PC model it takes 13 variables to account for the same amount of variation in F_1. Better illustration of WPC and PC differences in factor weights for $h = 0$ can be seen in Figure 1, which shows the cumulative sums of factor weights. While for PC model some variables do not get sufficient weights reflecting their importance in GDP forecasting, WPC may sometimes become too dependent on only a few variables.

Another important issue related to the big differences of WPC variable weights is that the inclusion of new variables might change extracted factors considerably. Due to this reason, the

[8] There is only one table for PC model's results as its variable weights do not change for different h in equation (10).

results become more dependent on a dataset used for forecasting. Despite the fact that in our case f_3 in equation (7) gives the best results in terms of the RMSE criterion, f_2 would be a safer choice for more conservative forecasting.

Figure 1: Comparison of factor weights' cumulative sums for WPC and PC



The differences in the WPC and PC objective functions are also reflected in Tables 5 – 8. The PC model seeks to retain as much information contained in X as possible and as a consequence it seems to be able to discriminate groups of variables, possessing similar variation patterns. Judging from the results in Table 5, factor F_1 could be interpreted as reflecting changes in consumer/industrial confidence, while F_2 reflects changes in industrial production, foreign trade. On the other hand, weights assigned by WPC should reflect the compromise of retaining as much X's information as possible and making sure that the information is relevant for GDP forecasting. However, results obtained using the WPC model might be somewhat hard to economically interpret and justify: e.g. goods transported by railways were the most important variable used to nowcast GDP, though it does not make a lot of sense economically. Also, we do not see the discriminated groups of variables in the WPC factors F_1 and F_2 – the same variables dominate in the structure of both factors.

Table 1: Average $\boldsymbol{R^2}$'s and p-values of coefficients for different $\boldsymbol{h}$ in equation (10).

		h=0	h=1	h=2
Average R^2	**PC**	0.28	0.12	0.11
	WPC	0.51	0.34	0.29
Average p-value of $\hat{\alpha}_{1,h}$	**PC**	0.233	0.320	0.48
	WPC	0.0002	0.045	0.044
Average p-value of $\hat{\alpha}_{2,h}$	**PC**	0.025	0.580	0.221
	WPC	0.507	0.352	0.351

Despite producing factors with less apparent interpretation, the WPC model's regression statistics are clearly superior to regression statistics of the PC model (average values of the regression statistics can be found in Table 1). This in turn reflects the gain of a different objective function used for the WPC model's factor extraction.

5.3 Small-scale factor model

The relative success of the WPC model for GDP nowcasting suggests that a small-scale factor model that encompasses only the most important and economically reasonable variables might be suitable for Lithuanian GDP nowcasting or forecasting. We included the following 5 variables to our small-scale factor model: narrow money aggregate, retail sales, industrial production (excluding construction), import and export.

As in the large-scale factor model, 2 factors were extracted to be used for GDP forecasting. We present the results of the RMSE and MAE criteria for the small-scale factor model in Table 4 of Appendix A.

Comparing RMSE and MAE criteria results, it seems that the small-scale factor model, which employs only 5 variables, outperforms the large-scale model. Hence, at least for the Lithuanian case, the quest of effectively incorporating much of the available data into the model to produce accurate short-term GDP growth forecasts remains a challenge.

6. Conclusions

We have performed a forecasting evaluation exercise for some factor models using the Lithuanian data. Contrary to the results of the previous study by Barhoumi et al. (2008), which also dealt with the Lithuanian data, we find that factor models outperform naïve benchmark forecasts. Different results of the two studies should be mainly attributed to longer time series available in our case.

The following three factor model specifications, popular in the forecasting literature, were tested in this forecasting evaluation exercise: PC, GPC and the state space model. The models produced rather similar results regarding their forecasting accuracy. It should be noted that none of the models were particularly useful when forecasting GDP two quarters ahead and were only slightly better than the benchmark model for one quarter ahead GDP forecasting.

To relate factors' extraction to their later use for GDP forecasting, the performance of several factor model modifications was also assessed. We find that the PLS model and factor models using variable preselection do not improve the forecasts for the Lithuanian data. The effect of the factor model modification with variables' weighting, depending on their absolute correlation with GDP growth, was studied in more detail. The results suggest that the WPC model might give better forecasts of the running quarter. On the other hand, WPC has several drawbacks: it uses a subjective weighting function, produces less interpretable factors and makes results more sensitive to changes in a variable set.

The small-scale factor model, using only 5 economically reasonable variables, actually produced more accurate forecasts than the large-scale model, which utilises 52 variables. This indicates that incorporating more data will not always result in the extraction of factors with better forecasting properties and that a small-scale factor model might be more applicable for the Lithuanian data.

References

Bai, J. and Ng, S. (2008). "Forecasting economic time series using targeted predictors". *Journal of Econometrics*, Elsevier, vol. 146(2), pages 304-317, October.

Bai, J. and Ng, S. (2002). "Determining the number of factors in approximate factor models". *Econometrica*, Econometric Society, vol. 70(1), pages 191-221, January.

Barhoumi, K., Benk S., Cristadoro R., Reijer A.D., Jakaitiene A., Jelonek P., Rua A., Rünstler G., Ruth K. and Van Nieuwenhuyze C. (2008). "Short-term forecasting of GDP using large monthly datasets - a pseudo real-time forecast evaluation exercise". Occasional Paper Series 84, European Central Bank.

Boivin, J. and Ng, S. (2006). "Are more data always better for factor analysis?“. *Journal of Econometrics*, Elsevier, vol. 132(1), pages 169-194, May.

Doz, C., Giannone, D., & Reichlin, L. (2011). "A two-step estimator for large approximate dynamic factor models based on Kalman filtering". *Journal of Econometrics*, Elsevier, vol. 164(1), pages 188-205, September.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). "Least angle regression". *Annals of statistics*, 32(2):407–499.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L., (2003). "The generalized dynamic factor model. One-sided estimation and forecasting". LEM Papers Series 2003/13, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Giannone, D., Reichlin, L. and Sala, L. (2005). "Monetary policy in real time". NBER Chapters, NBER Macroeconomics Annual 2004, Volume 19, pages 161-224 National Bureau of Economic Research, Inc.

Lee, T.H., Tu, Y. (2009), “Forecasting using supervised factor models”, working paper.

Rosipal, R., Krämer, N. (2006). “Overview and recent advances in partial least squares”. *Advances* 3940, 34-51.

Stock, J.H. and Watson, M.W., (2002). "Macroeconomic forecasting using diffusion indexes". *Journal of Business & Economic Statistics*, American Statistical Association, vol. 20(2), pages 147-62, April.

Thomaz, C.E., Gilson, A. G., Joaquim, F. P. Da Costa, Duncan, F. G., (2010). “A simple and efficient supervised method for spatially weighted PCA in face image analysis”, technical report.

Wold, H. (1966). “Estimation of principle components and related models by iterative least squares”. Krishnaiaah (ed) *Multivariate analysis*. Academic Press, New York.

Zou, H. and Hastie, T. (2005). "Regularization and variable selection via the elastic net". *Journal Of The Royal Statistical Society Series B*, Royal Statistical Society, vol. 67(2), pages 301-320.

Appendix A: Results of RMSE and MAE criteria

Table 2: Relative RMSE

Quarter forecasted			Preceding	Current			1 quarter ahead			2 quarters ahead		
Month of current quarter when forecasts were made			1	3	2	1	3	2	1	3	2	1
Model specification	PC	Unsupervised	0.74	**0.76**	0.85	0.88	**0.92**	**0.95**	**0.96**	**0.98**	**0.99**	0.92
		Variable selection	0.75	0.80	0.87	0.92	0.97	0.94	1.02	1.05	1.02	**0.91**
		WPC	**0.70**	0.76	**0.80**	0.93	0.96	1.01	0.98	1.02	**0.99**	0.97
	GPC	Unsupervised	0.76	0.77	0.84	0.95	0.97	0.98	0.98	0.99	1.00	0.93
		Variable selection	0.77	0.81	0.89	1.01	1.06	1.01	1.02	1.06	1.05	1.05
		WPC	0.71	0.76	0.81	0.93	0.95	0.99	0.99	1.03	1.02	0.99
	State space	Unsupervised	0.77	0.79	0.85	**0.82**	**0.92**	0.95	**0.96**	0.99	**0.99**	0.98
		Variable selection	0.74	0.79	0.84	0.91	0.97	**0.93**	0.97	0.99	**0.99**	0.99
		WPC	0.76	0.82	0.89	0.86	0.95	0.97	0.99	1.01	1.00	1.00
	PLS		0.88	0.95	0.92	0.94	0.97	0.97	1.10	1.05	1.05	1.03
Benchmark's MAE (%)			2.55	2.55	2.55	2.57	2.57	2.57	2.58	2.58	2.58	2.59

Table 3: Relative MAE

Quarter forecasted			Preceding	Current			1 quarter ahead			2 quarters ahead		
Month of current quarter when forecasts were made			1	3	2	1	3	2	1	3	2	1
Model specification	PC	Unsupervised	0.81	**0.82**	0.86	0.97	0.99	0.98	**0.93**	**0.93**	**0.96**	1.01
		Variable selection	0.79	0.86	0.85	1.11	1.13	1.04	1.05	1.11	1.08	1.04
		WPC	**0.78**	0.87	**0.84**	1.11	1.09	1.06	0.99	1.02	0.99	1.20
	GPC	Unsupervised	0.82	0.84	0.90	0.99	1.02	1.01	0.98	0.97	0.99	0.97
		Variable selection	0.86	0.87	**0.84**	1.09	1.25	1.08	1.12	1.10	1.11	1.07
		WPC	**0.78**	0.86	**0.84**	1.06	1.07	1.05	1.03	1.03	1.03	1.14
	State space	Unsupervised	0.82	0.86	0.91	**0.90**	**0.96**	**0.94**	**0.93**	0.96	**0.96**	**0.92**
		Variable selection	0.81	0.90	0.92	0.91	1.04	0.95	0.99	0.98	0.98	1.01
		WPC	0.82	0.90	0.92	0.93	0.98	1.01	0.97	1.00	0.97	0.97
	PLS		1.02	1.11	1.00	0.99	1.02	1.02	1.09	1.05	1.01	1.06
Benchmark's MAE (%)			1.47	1.47	1.47	1.47	1.47	1.47	1.48	1.48	1.48	1.50

Table 4: Relative RMSE's and MAE's of the small-scale factor model

Quarter forecasted		Preceding	Current			1 quarter ahead			2 quarters ahead		
Month of current quarter when forecasts were made		1	3	2	1	3	2	1	3	2	1
RMSE	Unsupervised PC	0.70	0.72	0.80	0.91	0.90	0.95	0.96	0.98	0.97	1.10
	WPC	0.65	0.71	0.81	0.92	0.89	0.95	0.97	0.99	0.97	1.04
MAE	Unsupervised PC	0.73	0.80	0.87	0.91	0.91	0.95	0.99	1.01	0.96	1.10
	WPC	0.66	0.78	0.83	0.95	0.92	0.95	1.00	1.03	0.96	1.15
Benchmark's RMSE (%)		2.55	2.55	2.55	2.57	2.57	2.57	2.58	2.58	2.58	2.59
Benchmark's MAE (%)		1.47	1.47	1.47	1.47	1.47	1.47	1.48	1.48	1.48	1.50

Appendix B: Graphs of PC and WPC forecasts[9]

Figure 2: Forecasts for h=0 months

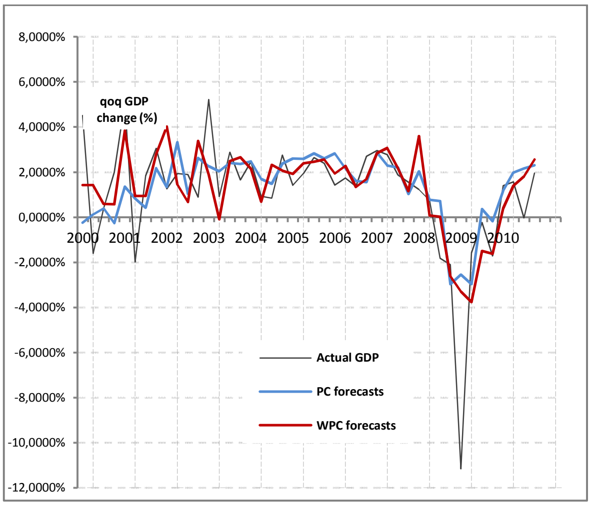


Figure 3: Forecasts for h=1 months

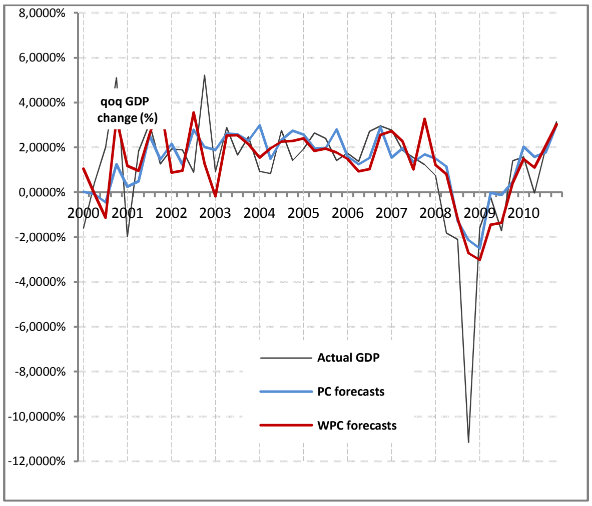


Figure 4: Forecasts for h=2 months

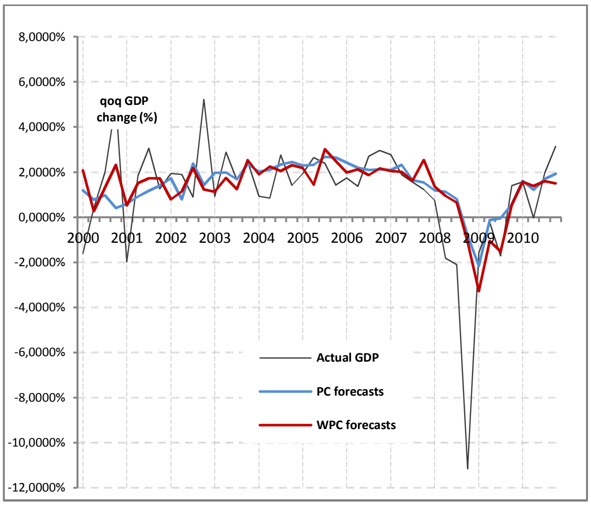
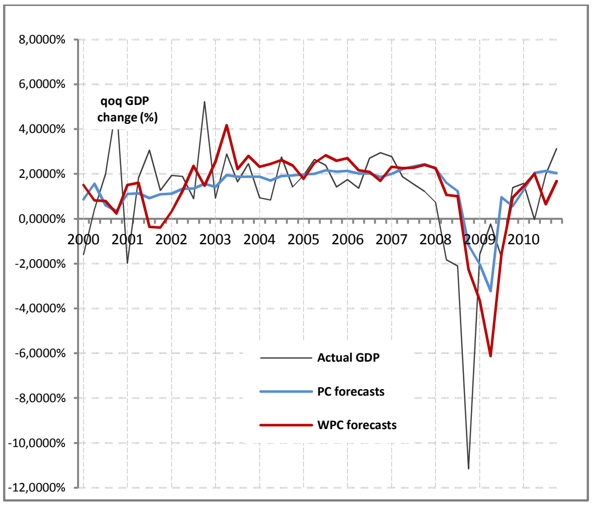


Figure 5: Forecasts for h=3 months



[9] Forecast horizons are defined as follows: horizon equal to 0 – forecasts of the previous quarter made in the first month of the current quarter, horizon equal to 1 – forecasts of the current quarter made in the third month of the current quarter, horizon equal to 2 – forecasts of the current quarter made in the second month of the current quarter and so forth.

Figure 6: Forecasts for h=4 months

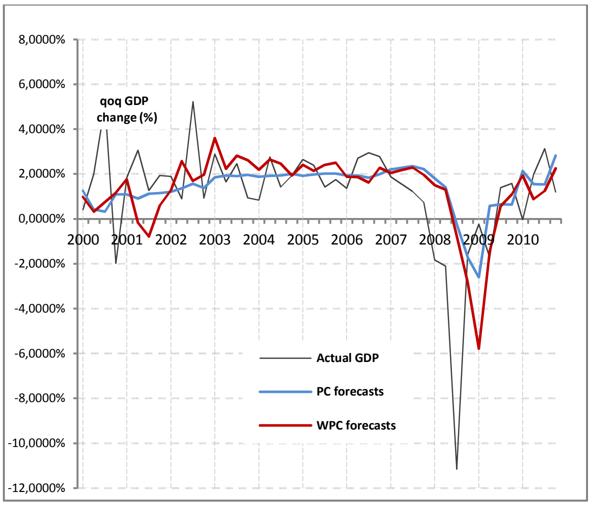


Figure 7: Forecasts for h=5 months

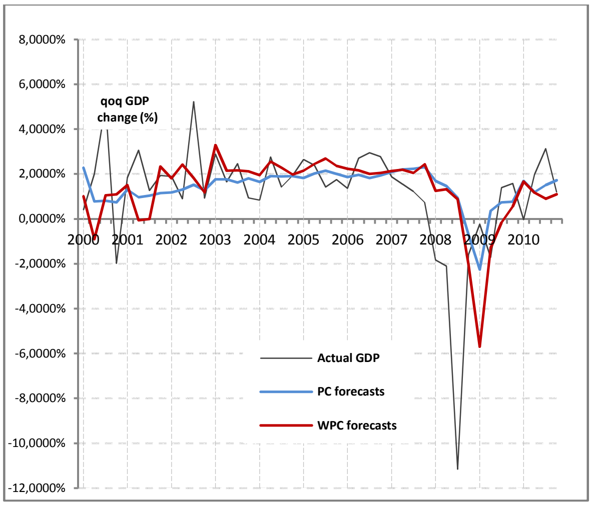


Figure 8: Forecasts for h=6 months

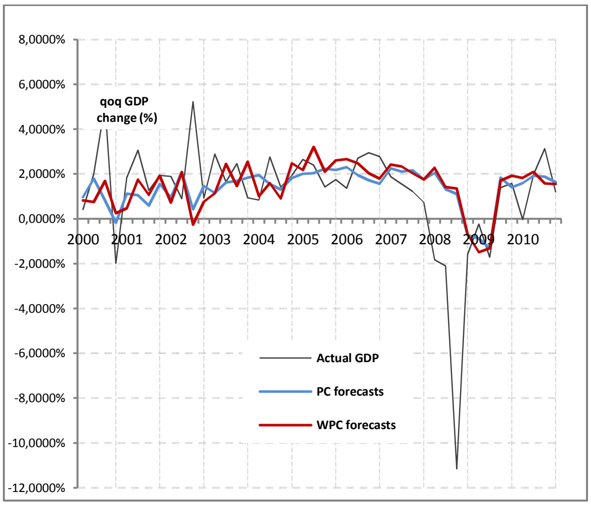


Figure 9: Forecasts for h=7 months

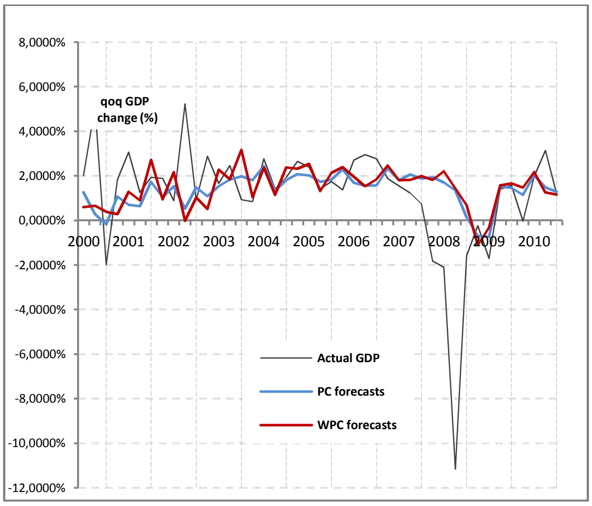


Figure 10: Forecasts for h=8 months

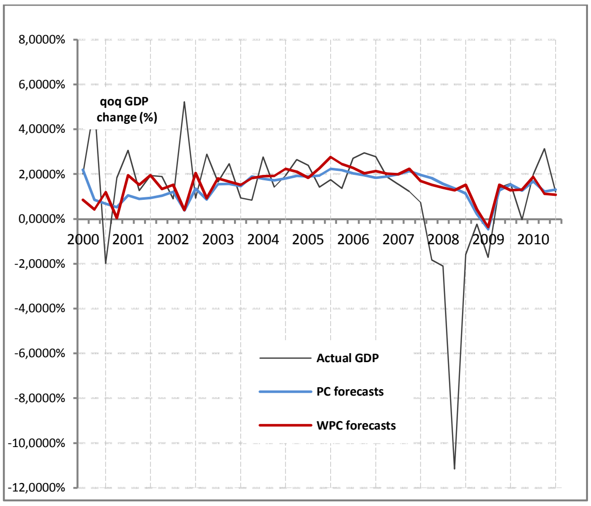
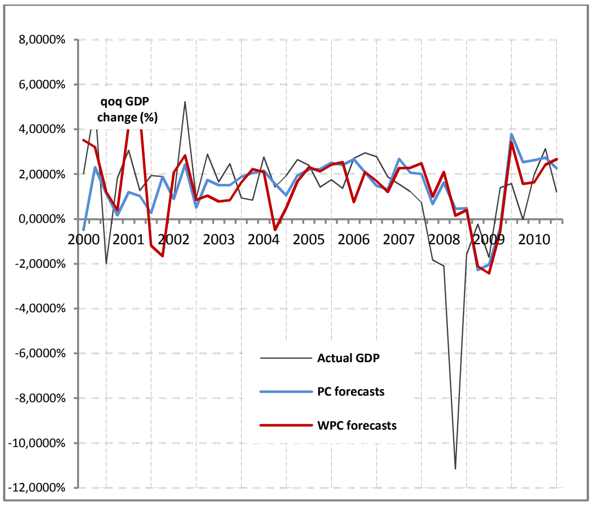


Figure 11: Forecasts for h=9 months



Appendix C: Comparison of variable weights assigned by WPC and PC

Table 5: Importance[10] of variables for PC factors $\boldsymbol{F_1}$ and $\boldsymbol{F_2}$ with $\boldsymbol{h = 0, 1, 2}$ in eq. (10)

	Factor F_1			**Factor F_2**	
	Variable	**Importance in the factor**		**Variable**	**Importance in the factor**
1.	Retail trade assessment of economic conditions in the upcoming 2-3 months	4.96%		IP-Manufacturing, excluding refined petroleum products	5.35%
2.	Construction survey: assessment of order books	4.72%		Import, all	5.26%
3.	Construction survey: employment expectations for the months ahead	4.60%		IP-Total Industry (excluding construction)	5.04%
4.	Industry survey: employment expectations	4.51%		Import of final goods	4.96%
5.	Retail trade survey: assessment of economic conditions in the preceding 2-3 months	4.44%		Export, all	4.39%

Table 6: Importance of variables for WPC factors $\boldsymbol{F_1}$ and $\boldsymbol{F_2}$ with $\boldsymbol{h = 0}$ in eq. (10)

	Factor F_1			**Factor F_2**	
	Variable	**Importance in the factor**		**Variable**	**Importance in the factor**
1.	Goods transported by railways	26.96%		Retail sales	17.35%
2.	Retail sales	13.55%		Goods transported by railways	17.30%
3.	IP-Total Industry (excluding construction)	8.65%		IP-Total Industry (excluding construction)	6.42%
4.	M1	4.40%		M1	4.65%
5.	Export, all	2.60%		Industry Survey: assessment of order-book levels	3.10%

[10] Importance of a standardized variable x_i in a factor F_k is defined as: $\widetilde{w}_{i,k} = \frac{|w_{i,k}|}{\sum_{i=1}^{n}|w_{i,k}|} \cdot 100\%$

Table 7: Importance of variables for WPC factors $\boldsymbol{F_1}$ and $\boldsymbol{F_2}$ with $\boldsymbol{h = 1}$ in eq. (10)

	Factor F_1			**Factor F_2**	
	Variable	**Importance in the factor**		**Variable**	**Importance in the factor**
1.	Industry Survey: Production expectations for the months ahead	21,69%		Import from Germany	10,55%
2.	Industry Survey: Production trend observed in recent months	8,19%		IP-Manufacture of furniture	8,13%
3.	Industry Survey: Employment expectations for the months ahead	5,90%		Industry Survey: Production expectations for the months ahead	7,49%
4.	M1	3,83%		Industry Survey: Production trend observed in recent months	5,16%
5.	Industry Survey: Assessment of export order-book levels	3,81%		Import from CIS countries	4,88%

Table 8: Importance of variables for WPC factors $\boldsymbol{F_1}$ and $\boldsymbol{F_2}$ with $\boldsymbol{h = 2}$ in eq. (10)

	Factor F_1			**Factor F_2**	
	Variable	**Importance in the factor**		**Variable**	**Importance in the factor**
1.	Export to CIS countries	20.40%		Export to CIS countries	11.11%
2.	Import from Poland	6.11%		Import from Poland	5.40%
3.	Industry Survey: Assessment of order-book levels	5.02%		OECD Composite leading indicator for Russia	5.20%
4.	OECD Composite leading indicator for Russia	4.32%		Industry Survey: Assessment of order-book levels	4.25%
5.	M2	3.59%		Retail trade survey: assessment of stock level	3.21%

Appendix D: Monthly dataset used for forecasting

Table 9: Monthly variables used in the study

Serial	Type	Description	Transformation
1	Survey	Construction survey: assessment of order books	-
2	Survey	Construction survey: employment expectations for the months ahead	-
3	Survey	Construction survey: trend of activity compared with preceding months	-
4	Survey	Construction survey: production’s price expectations	-
5	Survey	Euro area retail confidence indicator	-
6	Survey	Euro area consumer confidence indicator	-
7	Survey	Euro area industrial confidence indicator	-
8	Survey	Industry survey: assessment of order-book levels	-
9	Survey	Industry survey: assessment of export order-book levels	-
10	Survey	Industry survey: assessment of stock level	-
11	Survey	Industry survey: employment expectations for the months ahead	-
12	Survey	Industry survey: production trend observed in recent months	-
13	Survey	Industry survey: production demand expectations for the months ahead	-
14	Survey	Industry survey: selling price expectations for the months ahead	-
15	Survey	Retail trade survey: assessment of economic conditions in the preceding 2-3 months	-
16	Survey	Retail trade survey: assessment of economic conditions in the upcoming 2-3 months	-
17	Survey	Retail trade survey: assessment of stock level	-
18	Survey	Retail trade survey: employment expectations	-
19	Survey	Retail trade survey: orders placed with suppliers	-
20	Production	IP-Manufacture of chemicals and chemical products	(1-L)log
21	Production	IP-Manufacture of food products and beverages	(1-L)log
22	Production	IP-Manufacture of furniture	(1-L)log
23	Production	IP-Manufacture of textiles	(1-L)log
24	Production	IP-Manufacturing, excluding refined petroleum products	(1-L)log
25	Production	IP-Total Industry, excluding construction	(1-L)log
26	Trade	Lithuania trade with CIS, import	(1-L)log
27	Trade	Lithuania trade with CIS, export	(1-L)log
28	Trade	Lithuania trade with Germany, import	(1-L)log
29	Trade	Lithuania trade with Latvia, export	(1-L)log
30	Trade	Lithuania trade with Poland, import	(1-L)log
31	Trade	Import, all	(1-L)log
32	Trade	Import of final goods	(1-L)log
33	Trade	Import of investment goods	(1-L)log
34	Trade	Export, all	(1-L)log
35	Prices	HICP - Energy	(1-L)log

36	Prices	HICP - Total	(1-L)log
37	Prices	Retail prices	(1-L)log
38	Prices	PPI-Manufacture of chemicals and chemical products	(1-L)log
39	Prices	PPI-Total industry, excluding construction	(1-L)log
40	Financial	Dow Jones	(1-L)log
41	Financial	EUROSTOXX 50	(1-L)log
42	Financial	M1	(1-L)log
43	Financial	M2	(1-L)log
44	Financial	S&P 500	(1-L)log
45	Financial	Loans to households	(1-L)log
46	Other	Goods transported by railways	(1-L)log
47	Other	Klaipeda State Seaport and Butinge Sea Terminal, total goods handled	(1-L)log
48	Other	OECD Composite leading indicator for Germany	-
49	Other	OECD Composite leading indicator for Russia	-
50	Other	OECD Composite leading indicator for USA	-
51	Other	Registered unemployment rate	(1-L)
52	Other	Retail trade volume, except of motor vehicles and motorcycles	(1-L)log

List of Bank of Lithuania Working Papers

No 13: "Generating short-term forecasts of the Lithuanian GDP using factor models" by Julius Stakėnas, 2012.

No 12: "Profit Dynamics Across the Largest Euro Area Countries and Sectors" by Laurent Maurin, Moreno Roma and Igor Vetlov, 2011.

No 11: "Price Setting in Lithuania: More Evidence from the Survey of Firms" by Ernestas Virbickas, 2011.

No 10: "What Caused the Recent Boom-and-Bust Cycle In Lithuania? Evidence from a Macromodel with the Financial Sector" by Tomas Ramanauskas, 2011.

No 9: "Potential Output in DSGE Models" by Igor Vetlov, Tibor Hlédik, Magnus Jonsson, Henrik Kucsera and Massimiliano Pisani, 2011.

No 8: "The Implementation of Scenarios Using DSGE Models" by Igor Vetlov, Ricardo Mourinho Félix, Laure Frey, Tibor Hlédik, Zoltán Jakab, Niki Papadopoulou, Lukas Reiss and Martin Schneider, 2010.

No 7: "Wage and Price Setting Behaviour of Lithuanian Firms" by Ernestas Virbickas, 2010.

No 6: "Building an Artificial Stock Market Populated by Reinforcement-Learning Agents" by Tomas Ramanauskas and Aleksandras Vytautas Rutkauskas, 2009.

No 5: "Estimation of the Euro Area Output Gap Using the NAWM" by Günter Coenen, Frank Smets and Igor Vetlov, 2009.

No 4: "The Effects of Fiscal Instruments on the Economy of Lithuania" by Sigitas Karpavičius, 2009.

No 3: "Agent-Based Financial Modelling: A Promising Alternative to the Standard Representative-Agent Approach" by Tomas Ramanauskas, 2009.

No 2: "Personal Income Tax Reform: Macroeconomic and Welfare Implications" by Sigitas Karpavičius and Igor Vetlov, 2008.

No 1: "Short-Term Forecasting of GDP Using Large Monthly Datasets: A Presudo Real-Time Forecast Evaluation Exercise" by G. Rünstler, K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitienė, P. Jelonek, A. Rua, K. Ruth, C. Van Nieuwenhuyze, 2008.