European University Institute

Department of Economics

EUI Working Papers

ECO 2008/17

Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change

Anindya Banerjee Massimiliano Marcellino Igor Masten

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change

ANINDYA BANERJEE, MASSIMILIANO MARCELLINO

and

IGOR MASTEN

EUI Working Paper **ECO** 2008/17

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy

http://www.eui.eu/
http://cadmus.eui.eu/

Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change*

Anindya Banerjee
Department of Economics
European University Institute
Via della Piazzuola, 43
50133 Firenze (FI)
anindya.banerjee@eui.eu

Massimiliano Marcellino
IEP-Bocconi University, IGIER
and CEPR
Via Salasco, 5
20136, Milano
massimiliano.marcellino@uni-bocconi.it

Igor Masten
Department of Economics
University of Ljubljana
Kardeljeva pl. 17
1000 Ljubljana
igor.masten@ef.uni-lj.si

December 2006

Abstract

We conduct a detailed simulation study of the forecasting performance of diffusion index-based methods in short samples with structural change. We consider several data generation processes, to mimic different types of structural change, and compare the relative forecasting performance of factor models and more traditional time series methods. We find that changes in the loading structure of the factors into the variables of interest are extremely important in determining the performance of factor models. We complement the analysis with an empirical evaluation of forecasts for the key macroeconomic variables of the Euro area and Slovenia, for which relatively short samples are officially available and structural changes are likely. The results are coherent with the findings of the simulation exercise, and confirm the relatively good performance of factor-based forecasts also in short samples with structural change.

JEL Classification: C53, C32, E37
Keywords: Factor models, forecasts, time series models, structural change, short samples, parameter uncertainty

---

* The authors wish to thank Michael McCracken, David Rapach, Jack Strauss, Mark Wohar, Domenico Giannone, an anonymous referee and participants in the conference at St Louis University for their comments. Financial support from the Research Council of the European University Institute and the Italian MIUR is also gratefully acknowledged. Responsibility for any errors rests with us.

1. Introduction

Diffusion indexes extracted from dynamic factor models have been applied successfully in a number of papers to forecast macroeconomic variables. These include, among others, Stock and Watson (1999, 2002a, 2002b) for the US, Marcellino, Stock and Watson (2003) for the eleven countries originally in the Euro area, Artis, Banerjee and Marcellino (2005) for the UK, Schumacher (2006) for Germany, Bruneau, de Bandt, Flageollet and Michaux (2006) for France, and den Reijer (2006) for The Netherlands. The primary justification for the use of factor models in large datasets (where the number of variables N may exceed the sample size T) is their usefulness as a particularly efficient means of extracting information from many time series. This methodology also permits the incorporation of data at different vintages, frequencies, and time spans, thereby providing a clearly specified and statistically rigorous but economical framework for the use of large datasets in econometric analyses.

An interesting application of the dynamic factor model in a short sample context has been for forecasting the key macroeconomic indicators, e.g., GDP growth, inflation and interest rates, of the ten new members of the European Union,. Due to the period of transition, only short spans of time series are available for each of these countries, and parameter changes are likely However, despite these constraints, a large number of macroeconomic series of potential use in forecasting (for a given time span) are available for each country, and diffusion index-based forecasts can therefore be constructed. Banerjee, Marcellino and Masten (2006), using quarterly observations for the sample 1994:1-2002:2, show how diffusion index forecasts for these countries are often better than forecasts obtained from simple time series models, which is the alternative set of forecasting tools in this short-T context because of their parsimonious specification. [1]

The Euro area represents another interesting example of a short-T large-N forecasting context. This currency area has been in existence for only a short period of time, so that policymakers need to rely on limited spans of data and viable forecasting tools to conduct forward-looking policy. Furthermore, use of data from the pre-euro period has to account for the fact that the inauguration of the Euro area in 1999 and the introduction of the single currency in 2001 marked major shifts in policy for all the constituent countries. As shown by

[1] The time spans in the papers cited in the first paragraph are generally considerably longer. For example, the dataset in Stock and Watson (2002b) consists of monthly observations for the period 1959:1 to 1998:12, while Marcellino et al. (2003) use monthly and quarterly observations for the period 1982-1997. The monthly dataset in Artis et al. runs from 1972:1 to 1998:12.

Banerjee, Marcellino and Masten (2005), reliable leading indicators are difficult to find in such circumstances. However, diffusion index based forecasts still perform reasonably well.[2]

In this paper we examine in closer detail the reasons underlying the good performance of factor forecasts in the short sample context. We start by discussing briefly the key aspects of the competing modelling and forecasting approaches in Section 2. Section 3 presents the results of an extensive Monte Carlo analysis of the performance of factor-based forecasts in short samples, possibly subject to parameter changes. Section 4 illustrates the issues empirically, by comparing alternative forecasting methods for the key macroeconomic variables of the Euro area and Slovenia. The latter is a newly acceded country to the European Union, the first among the new member states to adopt the euro (in 2007). Transitional changes in this country include not only the switch from a planned to a free-market economy but also the changes involved in adopting the euro. Therefore, in light of the discussion above, forecasting the developments in Slovenian macroeconomic variables represents a stern test of the efficacy of the various methods considered. Section 5 summarizes and concludes the paper.

2. Methodology

This section, which is based on Banerjee et al. (2005, 2006), reviews the competing forecasting approaches we consider both in the Monte Carlo analysis in Section 3 below and, more particularly, in the empirical analysis discussed in Section 4. We also state the criteria used to evaluate the relative merits of the alternative forecasts, see *e.g.* Marcellino et al. (2003) or Artis et al. (2005) for additional details.

All forecasting models are specified and estimated as a linear projection of an h-step-ahead variable, y^h_{t+h}, onto t-dated predictors, which at a minimum include lagged transformed values (denoted y_t) of x_t, the series of interest. More precisely, the forecasting models all have the form,

$$y^h_{t+h} = \mu + \alpha(L)y_t + \beta(L)'Z_t + \varepsilon^h_{t+h} \quad (1)$$

where $\alpha(L)$ is a scalar lag polynomial, $\beta(L)$ is a vector lag polynomial, μ is a constant, and Z_t is a vector of predictor variables. Marcellino, Stock and Watson (2006) present a

[2] Two other examples of the usefulness of factor models with fewer than 50 time-series observations are Matheson (2005) for the case of New Zealand and Breitung and Eickmeier (2005) who provide an example of the use of dynamic factor models in macroeconomic analysis for the case of the Euro area using fewer than 45 quarterly observations of macroeconomic data.

comparison of this *h*-step projection method with the more standard approach of specifying a model for y_t and then solving it forward to obtain a forecast for y_{t+h}. However, due to the short sample available, both in the Monte Carlo evaluation of Section 3 and in the empirical analysis of Section 4, we focus on one-step ahead forecasts, so that $h = 1$ in (1). In the empirical analysis monthly data is used for the case of the Euro area, while the frequency is quarterly for Slovenia,

The construction of y^h_{t+h} depends on whether the series is modelled as I(0), I(1) or I(2), where series integrated of order *d*, denoted I(*d*), are those for which the *d-th* difference (Δ^d) is stationary. Indicating by *x* the series of interest (usually in logarithms), in the I(0) case, $y^h_{t+h} = x_{t+h}$ and $y_t = x_t$. In the I(1) case, $y^h_{t+h} = \sum_{t+1}^{t+h} \Delta x_s$ so that $y^h_{t+h} = x_{t+h} - x_t$, while $y_t = x_t - x_{t-1}$. In words, the forecasts are for the growth in the series *x* between time period *t* and *t*+*h*. Finally, in the I(2) case, $y^h_{t+h} = \sum_{t+1}^{t+h} \Delta x_s - h\Delta x_t$ or $y^h_{t+h} = x_{t+h} - x_t - h\Delta x_t$, *i.e.*, the difference of *x* between time periods *t* and *t*+*h* and *h* times its growth between periods *t-1* and *t*, and $y_t = \Delta^2 x_t$. This is a convenient formulation because, given that x_t and its lags are known when forecasting, the unknown component of y^h_{t+h} conditional on the available information is equal to x_{t+h} independently of the choice of the order of integration. This makes the mean square forecast error (MSE) from models for second-differenced variables directly comparable with, for example, that from models for first differences only. The MSE is computed as the average of the sum of squares of all the comparisons between the actual value of the variable and its forecast (under any of the methods given in Section 2.1 below).

2.1 Forecasting models

The various forecasting models we compare differ in their choice of Z_t in equation (1). Let us list the forecasting models and briefly discuss their main characteristics.

Autoregressive forecast (ar_bic). Our benchmark forecast is a univariate autoregressive (AR) forecast based on (1) excluding Z_t. In common with the literature, we choose the lag length using an information criterion, the BIC, starting with a maximum of 6 lags. While this model is very simple, the resulting forecasts are typically rather accurate, see e.g. Marcellino (2006).

Autoregressive forecast with second differencing (ar_bic_i2). Clements and Hendry (1999) showed that second differencing the variable of interest improves the forecasting performance of autoregressive models in the presence of structural breaks. This is an interesting option to be considered in the case of most of the new EU member states, which have undergone several economic and institutional changes even after the fairly rapid transition to a market economy. This model corresponds to (1), excluding Z_t and treating the variable of interest as I(2).

Autoregressive forecast with intercept correction (ar_bic_ic). An alternative remedy in the presence of structural breaks over the forecasting period is to put the forecast back on track by adding past forecast errors to the forecast, *e.g.* Clements and Hendry (1999) and Artis and Marcellino (2001). They showed the usefulness of the simple addition of the *h*-step ahead forecast error. Hence, the forecast is given by $\hat{y}_{t+h}^{h} + \varepsilon_t^h$, where $\hat{y}_{t+h}^{h}$ is the ar_bic forecast and ε_t^h is the forecast error made when forecasting y_t in period *t-h*. Since both second differencing and intercept correction increase the MSE when not needed, by adding a moving average component to the forecast error, they are not costless and should only be used if needed. However, the empirical applications we consider are such that macroeconomic series are very likely to have breaks due to policy changes, implying that second differencing and intercept correction are options well worth considering.[3]

VAR forecasts (varf). Vector autoregressive (VAR) forecasts are constructed using equation (1) with chosen regressors Z_t. In particular, in the empirical analysis in Section 4, Z_t includes lags of GDP growth, inflation, and a short-term interest rate. Intercept corrected versions of the forecasts are also computed (varf_ic).

Factor-based forecasts. These forecasts are based on setting Z_t in (1) to be the estimated factors from a dynamic factor model, the so-called diffusion indexes, along the lines of Stock and Watson (2002b), to which we refer for addition details. While other methods are available for factor extraction, see e.g. Forni, Lippi, Hallin and Reichlin (2000, 2005) and Kapetanios and Marcellino (2006), or for forecasting in the presence of many predictors, see e.g. the review in Stock and Watson (2006), Stock and Watson's (2002b) approach performed well in a variety of empirical forecasting applications.

Under some technical assumptions (restrictions on moments and stationarity conditions), the column space spanned by the dynamic factors f_t can be estimated consistently by the principal components of the $T \times T$ covariance matrix of the X's. The factors can be considered as an exhaustive summary of the information contained in a large dataset.

[3] In the case of the Euro area, one main candidate for a break to consider is the introduction of the Euro in 1999. Slovenia, on the other hand, entered the ERM II system in 2004.

It is also worth mentioning that the principal-component-based factor estimate remains consistent even in the presence of limited time variation in the parameters of the underlying factor model. Such a property can be very convenient for analyzing the economies of the new European member states, which are under constant evolution, or more generally time series possibly subject to changes of regime. However, the effects of time variation in short samples also deserve careful analysis, which is undertaken in Section 3.

In the empirical applications, we primarily consider three different factor-based forecasts. First, in addition to the current and lagged y_t up to 4 factors and 3 lags of each of these factors are included in the model (fdiarlag_bic).[4] Second, up to 12 factors are included, but not their lags (fdiar_bic). Third, up to 12 factors appear as regressors in (1), but no current or lagged y_t is included (fdi_bic). For each of these three classes of factor-based forecasts the model selection is based on BIC. The factors can be extracted from the unbalanced panel of available time series (prefix fac), or from the balanced panel (prefix fbp) and we consider them both. The former contains more variables than the latter, and therefore more information. The drawback is that missing observations have to be estimated in a first stage, which could introduce noise in the factor estimation (see Angelini, Henry and Marcellino, 2006).

In order to evaluate the forecasting role of each factor, for the unbalanced panel, we also consider forecasts using a fixed number of factors, from 1 to 6 (fdiar_01 to fdiar_06 and fdi_01 to fdi_06).

Intercept-corrected versions of all the diffusion index based forecasts are also considered.

Finally, we construct a pooled factor forecast by taking a simple average of all the factor-based forecasts. Pooling is done separately over factor models without intercept correction (denoted f_pooled) and with intercept corrections (denoted f_ic_pooled). The pooled factor forecasts have particular informative value. In fact, since we consider many different versions of factor models, it should not be surprising to find at least one model that forecasts better than simple linear models. The average performance of factor models in this respect tells us whether factor models are in general a better forecasting device or if their relative good performance is limited only to some special sub-models.

[4] The notation "fdi"(forecast model_diffusion_index) derives from the work of Stock and Watson (1998), Marcellino et al. (2003) and Artis et al. (2005) inter alia and generally denotes a forecast model based on diffusion indexes (estimated factors).

2.2 Forecast Comparison

The forecast comparison is conducted in a simulated out-of-sample framework where all statistical calculations are done using a fully recursive methodology. In the empirical examples, the models are first estimated on an initial data span, for example from 1994:1 to 2000:2, and 1-step-ahead forecasts are then computed. The estimation sample is then augmented by one quarter and the corresponding 1-step-ahead forecast is computed. Every quarter, (*i.e.* for every augmentation of the sample) all model estimation, standardization of the data, calculation of the estimated factors, etc., is repeated until the end of the data span.

The forecasting performance of the various methods described is examined by comparing their simulated out-of-sample MSE relative to the benchmark autoregressive (AR) forecast (ar_bic). West (1996) standard errors are computed around the relative MSE in the empirical analysis of Section 4.[5]

3. Monte Carlo Experiments

As noted previously, almost all existing examples of applications of factor forecast methods have relied on datasets where the time dimension is very (or at least fairly) long. By contrast, we are interested here in cases where T is small, since in many interesting macroeconomic panels (such as those for the transition countries or the Euro area after its creation) T rarely exceeds 30 observations.

Though in theory the time dimension is not a problem, as long as the longitudinal dimension of the dataset to be used for factor extraction is large enough, in practice the feasibility and relevance of factor forecasts can be questioned for such a short sample. Therefore, in this section we compare the performance of AR and factor forecasts by means of simulation experiments, using datasets with a short time span.

Using artificially generated data, our attempt is to understand the sensitivity of the performance of factor- and non-factor methods to sample size T and longitudinal dimension N. We also explore the sensitivity of such methods to various other features likely to characterize the data – such as the degree of stationarity of the factors, the amount of autocorrelation and the presence of structural change. The latter set of experiments has a great deal of relevance for practical applications, but the impact of these features of the processes generating the data on the performance of factor forecasts has not been, to the best of our knowledge, studied in any detail in the literature.

[5]West (1996) standard errors are first computed for the MSE of the benchmark model and of the competing model. The standard errors of the ratio of the MSEs are then computed using the delta method.

In the first subsection, we describe the design of the experiments. In the second subsection, we discuss the results. In the third subsection, we provide an explanation for the good performance of AR forecasts in some cases. In the final subsection, we summarize the main findings.

3.1. Design of experiments

The Monte Carlo design is taken from Stock and Watson (1998) and adapted for the purposes of this paper. The data are generated by a dynamic factor model that allows for autoregressive factors, auto and cross-correlation in idiosyncratic errors and time-varying parameters. A balanced panel of data is generated as follows:

$$x_{it} = \lambda_{it}^{'} f_t + e_{it} \tag{2}$$

$$\lambda_{it} = \lambda_{it-1} + (c/T)\zeta_{it} \tag{3}$$

$$f_t = A_t f_{t-1} + u_t, A = \alpha_t I_r \tag{4}$$

$$\alpha_t = d(\alpha_{t-1} + 1/T\,\eta_t) + (1-d)\alpha_1 I(T_B) + (1-d)\alpha_0,\ \alpha_0 = \overline{\alpha} \tag{5}$$

$$I(T_B) = \begin{cases} 0, \forall t \le T_B \\ 1, \forall t > T_B \end{cases} \qquad d = \begin{cases} 0, \text{breaking} & \alpha \\ 1, \text{time varying} & \alpha \end{cases} \tag{5a}$$

$$(1-aL)e_{it} = (1+b^2)v_{it} + bv_{i+1,t} + bv_{i-1,t} \tag{6}$$

$$y_t = \iota' f_{t-1} + \varepsilon_t \tag{7}$$

where $i = 1,\ldots,N$ and $t = 1,\ldots,T$. In equation (2), the common factors are indicated with f_t, their associated loadings with λ_{it}, and the number of factors, r, varies from 1 to 5. The variables e_{it} in (2), which represents the idiosyncratic component for x_{it}, v_{it} in (6), η_t in (5) and ε_t in (7) are i.i.d. $N(0,1)$, while ζ_{it} in (3) and u_t in (4) are i.i.d. $N(0,I_r)$. The error terms in u_t are independent of e_{it}, v_{it}, η_t, ε_t and ζ_{it}. Each variable x_{it} is standardized prior to estimation. The scalar variable to be forecast is indicated by y_t, and its expected value coincides with the sum of the factors in the previous period, namely, ι in (7) is an $r\times 1$ vector of 1's.

The parameter α_t in (4) measures the persistence of the factor series, for which we consider three cases in order to analyze the impact of structural change. The first is the case of stable and fixed α ($d = 0$ and $T_B = 0$). The second is that of continuously time-varying persistence, parameterized by setting d to 1. The starting values for α_t, given by α_0, successively take on the values $\overline{\alpha} = \{0.3, 0.5, 0.7\}$. These are also the three values considered for α when it is fixed. A persistence parameter of 0.7 is taken as the maximum, which might seem at first sight too small to replicate persistence of factors sometimes found in the data. However, this restriction prevents, for all our choices of *T*, the time-varying persistence from persistently drifting above unity which happens when considering higher initial values for alpha. Note also that time-varying persistence of the type considered here can generate factors with very high persistence. The third is the case of a discrete break in persistence of factors. We let $d = 0$ and $T_B = T/2$, i.e. the break occurs in the middle of the sample. α_1 takes on two values: 0.4 when α_0 is set to 0.3, and -0.4 when α_0 is set to 0.7. In other words, we model two types of discrete breaks in the persistence of factors: in the first persistence increases from 0.3 to 0.7, while in the second it decreases from 0.7 to 0.3. In all three cases the factors are standardized with their (estimated) standard deviation to achieve unitary variance, and their loadings are kept fixed ($c = 0$ in equation (3)).

Next, in order to evaluate the role of a change in the variance of the factors, we double the variance of factors, while keeping their persistence unchanged. This implies that a higher weight in generating the overall variability of the variables in the panel is attributed to the variability of factors relative to the variability of the idiosyncratic component in equation (2).

Structural change in the panel can also be due to time-varying factor loadings. For these experiments, equation (3) becomes relevant. The time-varying parameter (TVP) c is thus set to either 0 (no TVP) or 5 (TVP) depending on whether or not time variation is modelled or not.[6]

We do not evaluate the consequences of changes of the factor parameters in the forecasting equation (7), even though this is an important practical source of forecast failure, since this case has been extensively studied in the literature, see e.g. Clements and Hendry (1999).

Finally, for the setting with correlated idiosyncratic errors, parameters a and b are set to 0.5 and 1 respectively. In terms of cross-sectional correlation this implies $E(e_{i,t}e_{i\pm1,t}) = 1/3$, $E(e_{i,t}e_{i\pm2,t}) = 1/6$ and zero otherwise.

[6] Note that the variance of factor loadings depends on c/T. This implies asymptotically constant factor loadings, which is needed for consistent factor estimation. In our Monte Carlo simulations this implies that with increasing T factor loadings change less than when T is small. The same holds also in the case where we consider similar time variation in the persistence of factors.

Four different configurations for (T, N) are considered, namely (T = 30, N = 50), (T = 50, N = 50), (T = 50, N = 100) and (T = 150, N = 50), as being representative of relevant panel sizes in the empirical examples studied.

The first configuration, T = 30, N = 50, is the benchmark and is relevant for forecasting with quarterly data in many transition economies (including the new EU members) and the Euro area following the introduction of the euro.

The second, T = 50, N = 50, considers a slightly longer time series, which should give us an impression of how forecasting circumstances may change with the passage of time, i.e. as more time-series observations become available (but the parameters of the model remain constant). While this increment of 20 observations may be small in absolute value, it still represents a large relative increase in the length of the panel.

The third configuration, T = 50, N = 100, considers the scenario of a time series with the T dimension corresponding to those above but the larger N allows for the evaluation of the relative merits of considering more variables in factor extraction. Such a scenario is highly relevant both for new EU members as they adopt EU standards in their data collection practices and also for the Euro area where more and more aggregate series become gradually available. While a larger longitudinal dimension improves the precision of the factor estimates when the additional variables are driven by the same factors, it can create serious problems when the added variables are driven by different factors, in particular if the latter have low correlation with the target variable in the forecasting exercise. For this reason Boivin and Ng (2006) have suggested pre-selecting the variables to be used for factor estimation, based on their correlation with the target variable. While this is not an issue in our Monte Carlo experiment, it will become relevant in the ensuing empirical applications.

The last configuration, T = 150, N = 50, particularly suits the empirical application detailed below of forecasting Euro area variables using monthly data also from the period before the monetary union was actually formed (i.e. from 1991 onwards). It also allows us to consider the effects of a major increase in the temporal dimension, from 30 to 150, keeping the number of variables fixed.

The factors are estimated by principal components as described previously. We focus on the comparison of factor forecasts with forecasts made by using AR models.

Three types of simple AR forecasts are produced: with fixed lags of 1 and 3 and with lag length chosen by BIC. We have included the parsimonious fixed lag length AR specification in the comparison since information criteria have an asymptotic justification and may not work well in short samples, while parsimony can be a plus in this context since it reduces estimation uncertainty.

The factor models used in the comparison are four in number. First, we generate forecasts using the known coefficients and factors from the data generation process (*fdi_dgp*). Second, we consider using the true factors but estimated coefficients to generate forecasts (*fdi known factors*). Third, we use the estimated factors and the estimated coefficients (*fdi fully estimated*), assuming that the true number of factors is known. Finally, we generate forecasts from a regression of y on own lags and on current and lagged values of estimated factors (*fdiarlag_bic*). The maximum lag length is 3 and up to r factors are included. Model selection is by BIC. [7]

The mean square error (MSE) of each model is computed relative to the AR(1) model. The numbers can of course easily be re-standardized in order to use the ar_bic model as the benchmark.

Table 1 summarizes the values of the key parameters of the data-generating process in equations (2) – (7) used in the different Monte Carlo experiments. All experiments are repeated for the number of factors, r, varying from 1 to 5. The number of Monte Carlo replications is 20000.

Table 1: Parameterization of the models in the Monte Carlo

DGP setting		***a***	***b***	***c***	***d***	T_B	α_0	α_1
Basic DGP		0	0	0	0	0	0.3, 0.5, 0.7	0
Time-varying alpha		0	0	0	1	0	0.3, 0.5, 0.7	0
Discrete break in alpha	0.3/0.7	0	0	0	0	*T/2*	0.3	0.4
	0.7/0.3	0	0	0	0	*T/2*	0.7	-0.4
Time-varying lambda		0	0	5	0	0	0.3, 0.5, 0.7	0
Correlated errors		0.5	1	0	0	0	0.3, 0.5, 0.7	0
Double variance		As with Basic DGP, but double variance of factors						

3.2 Results

Figures 1 to 5 summarize the results of the Monte Carlo simulations, and the titles of the figures reflect the classification of the experiments in Table 1. The key distinctions, among

[7] As we note below, intercept correction and double-differencing are generally not useful forecasting methods for our empirical examples, except for the case of Slovenian core inflation which may be due to a clear break in the inflation target that occurred after 2002, reflecting disinflation to meet the Maastricht criteria. Pooling of forecasts also does not appear to be an effective device for our datasets. These methods are thus excluded from the comparisons in our simulation experiments.

the DGPs considered, are made according to the number of factors (from 1 to 5) and their persistence. All the results are for h = 1 (one-step ahead forecasts).[8]

Figure 1 presents the results for different values of N and T and the basic DGP, where $d=a=b=c=T_B=0$ and $\overline{\alpha} = \{0.3, 0.5, 0.7\}$. As expected, the best forecasts are given by the empirically implausible model *fdi_dgp* where every feature of the data generation process is known. Comparing the forecasts generated by *fdi known factors* and *fdi fully estimated* with *fdi_dgp*, we find that the effects of estimation uncertainty can be quite marked, especially when T = 30, the number of factors increases towards 4 or 5 and the factor loadings need to be estimated. Notice, in particular, that while a larger number of factors improves the performance of *fdi_dgp*, it markedly worsens that of *fdi fully estimated.* [9]

A finding of note is the bad performance of BIC-based factor forecasts, in particular when the number of factors is large, while BIC-based AR forecasts are typically better than AR(3) forecasts but comparable to the AR(1) ones. These results occur since BIC penalizes extensive parameterizations and in such short samples it ends up selecting models with too few factors or with just one lag of the dependent variable.

An increase of the temporal dimension to 50 already improves the performance of the BIC-based factor forecasts, and more generally of the factor forecasts with estimated parameters. While the factors are already fairly accurately estimated with 30 temporal observations, the precision of the estimators of the parameters of the forecasting equation increases substantially.

Further increases in the sample size (either increasing the N dimension from 50 to 100 or the T dimension from 50 to 150) additionally improve the relative performance of factor models, which is expected because the DGP has a factor structure. With a larger value for N or T, the persistence parameter α no longer has an important effect on the performance of factor models. It is also worth adding that the forecasting performance of AR models relative to the benchmark appears largely unaffected by increases in T and N, although since factor models show improved performance with increases in size of the panel, AR models lose efficiency relative to factor models as the size of the panel increases.

Figures 2 and 3 repeat the comparison exercise by allowing for structural change. In particular, Figure 2 deals with the cases where either $d = 1$ (continuously time-varying α) or

[8] We have also computed results for a longer forecasting horizon, h = 4, and for several other configurations of T and N. These are available from us upon request.

[9] This occurs especially when the persistence of factors approaches unity. In particular, with a persistence parameter of 0.9, it can happen that the relative MSE of the factor model exceeds one when the number of factors considered in the DGP is large. These results are available upon request.

there are discrete breaks in the persistence parameter half way through the sample. The starting values α_0 for the former case are given by 0.3, 0.5 and 0.7 respectively. Figure 3 reports the results in the presence of changes in the factor loadings denoted by λ.

Figure 2 again shows the dramatic underperformance of factor models when T = 30 and the DGP is generated with high number of factors and BIC selection is used in estimation, although their performance again improves once T increases to 50 or more. The comparisons across the methods of generating forecasts are not much affected by allowing for structural change in α, a finding which is coherent with our previous observation that the persistence parameter does not significantly affect the relative performance of factor vs. non-factor methods.

As reported in Figure 3, the consequences of continuously time-varying factor loadings are rather severe when T = 30, the most empirically relevant size of T, from the point of view of the empirical examples in our paper,. Factor forecasts based on the BIC criterion can perform worse than the AR(1) benchmark when r = 4, 5. This deterioration is more serious when compared to those arising from previous cases where the loadings do not change. However, increasing T to 50 helps greatly.

The results for a higher variance of the factors are reported in Figure 4. The consequences are more substantial: with more of the variance in the dataset explained by the factors, factor methods are strongly dominant even with T = 30, although there are some deteriorations evident as the number of factors increases. The relative dominance of factors increases with larger values of T and/or N. Of course, a reduction in the variance of the factors has the opposite effect, with BIC-based forecasts often outperformed by the AR alternatives. This can be empirically relevant when forecasting variables with a substantial idiosyncratic component. When the variance of the factors changes in the middle of the sample, intermediate results are obtained with respect to those in Figures 1 and 4.

Finally, Figure 5 reports the results for correlated errors, over time and variables, which largely match the results reported for the basic DGP in Figure 1.

3.3 Explaining the good performance of AR forecasts

To provide some intuition for the good performance of the AR forecasts in some of the cases described above, in particular when the number of factors is large and the temporal dimension is small, let us assume that the-factor model generating the variable of interest is

$$y_{t+1} = \gamma f_t + \varepsilon_{t+1} \qquad (8)$$

where ε_t are i.i.d $(0, \sigma_\varepsilon^2)$, and each factor follows an AR(1) model with the same persistence parameter (as in the Monte Carlo DGP):

$$f_{it} = \alpha f_{it-1} + u_{it} \qquad (9)$$

where each u_{it} is i.i.d $(0, \sigma_{iu}^2)$ and the errors ε and u are assumed to be independent for all t and s. If α and γ are known and f_t is observable, then the one-step ahead forecast error (conditional on the past history of y_t observable at time t) is easily seen to be given by σ_ε^2.

The model in (8)-(9) implies that y_t can be written as an ARMA (1,1) process

$$y_{t+1} = \alpha y_t + z_{t+1} + \theta z_t$$

where $z_{t+1} = g(\varepsilon_{t+1}, u_t)$. The forecast error variance is now given by $\sigma_z^2 = E(z_{t+1}^2)$ and it may be shown that $\sigma_z^2 > \sigma_\varepsilon^2$, so that the ARMA forecast is less efficient than the factor forecast when the parameters and the factors are known.

Yet, in practice both the parameters of the model and the factor are unknown and must be estimated – the latter by extraction from a large dataset and the former by a regression of y_{t+1} on the estimated factor. Bai and Ng (2006) show that, even in more general models, estimation of the parameters adds $O(T^{-1})$ uncertainty to the forecast while estimation of the factor adds $O(N^{-1})$ uncertainty. In other words, the factor-based forecast error variance for the case where both the factor and the parameters of the model have to be estimated is given by $\sigma_\varepsilon^2 + O(T^{-1}) + O(N^{-1})$. An additional term should be added to this quantity when the specification of the forecasting model is not known and has to be selected, e.g., by BIC. The factor-based error variance can now be larger than its counterpart for the estimated ARMA forecast, even if the ARMA model is approximated by a finite order AR. These effects are important at smaller sample sizes but are muted by an increase in either T or N, so that the performance of factor models improves with such an increase, as borne out by the results.

3.4 Summary

In summary, the following points regarding the effects for forecasting of instability of the parameters of the factor model may be noted:

(a) continuous changes in factor persistence do not seem to matter, even in short samples;

(b) discrete changes do matter but the impact on relative performance of factor methods leads either to improvement or deterioration, depending upon the value of the (starting) persistence parameter, the direction of the change and the magnitude of the T and N dimensions;

(c) time varying factor loadings are important except when T and N are large – i.e. for the size of our panels, estimated below, time variation of the loadings are likely to be very important;

(d) taking all the results of the simulation exercise into account, the ranking of the impact of the different kinds of stability is (c) to (b) to (a);

(e) factor models outperform AR models in the majority of cases, even in short samples subject to changes.

Two additional findings that may be highly relevant for empirical analyses are:

(f) the number of factors in the underlying data generating process is quite important when the sample is short and model selection is BIC-based, since the relative forecast performance deteriorates when the number of factors is large;

(g) the variance of the idiosyncratic component of the target variable is important; and the forecasting performance of factor models is reduced when the former is large.

4. Two empirical examples

In this section we present the results of two empirical exercises to provide practical content to our simulation findings, and to use our results so far to judge the outcome of forecasting with factor models in small T panels.

In the first subsection, we describe the datasets for the Euro area and for Slovenia. In the second subsection we present the results of the forecast comparison exercise for the Euro area, and in the third subsection for Slovenia. In the final subsection, we summarize the main findings.

4.1. The data

The dataset for the Euro area contains 58 monthly series, spanning over the period 1991:2 – 2005:10 (N=58, T=177) and is collected from OECD Main Economic Indicators and Eurostat. The dataset for Slovenia contains 95 quarterly series for the period 1994:1 – 2005:4 (N=95, T=48). For Slovenia the sources are the National Statistical Office and the Bank of Slovenia.

The datasets broadly contain output variables (GDP components, industrial production and sales); labour market variables (employment, unemployment, wages); prices (consumer, producer); monetary aggregates; interest rates (different maturities, lending and deposit rates);

stock prices; exchange rates (effective and bilateral); imports, exports and net trade; survey data; and other miscellaneous series. A complete list of the variables is reported in the data appendix.

Following Marcellino et al. (2003), the data are pre-processed in three stages before being modelled with a factor representation. First, we pass all the series through a seasonal adjustment procedure as very few series are originally reported as seasonally adjusted. Seasonal adjustment is performed with the original X-11 ARIMA procedure. Second, the series are transformed to account for stochastic or deterministic trends, and logarithms are taken of all nonnegative series that are not already in rates or percentage units. We apply the same transformations to all variables of the same type. The main choice is whether prices and nominal variables are I(1) or I(2). The I(1) case is our baseline model and all the results reported below apply to this choice. Banerjee et al. (2005) have also recomputed all the results treating prices, wages, monetary aggregates and nominal exchange rates as I(2) variables, showing that there are no substantial changes in the ranking of the forecasts. Variables describing real economic activity are treated as I(1), whereas survey data are treated as I(0).[10] All series were further standardized to have sample mean zero and unit sample variance.

Finally, the transformed seasonally adjusted series are screened for large outliers (outliers exceeding six times the inter-quartile range). Each outlying observation is recoded as missing data, and the EM algorithm is used to estimate the factor model for the resulting unbalanced panel.

Among the available variables, we have chosen to report forecasting results for HICP inflation (energy excluded), manufacturing output growth and unemployment for the Euro area, and CPI inflation (all items), core inflation (energy and food prices excluded) and GDP growth for Slovenia. These are also the variables of central importance for policymakers. Note, however, that the generality of the approach would easily allow us to extend the analysis to other variables of interest.

4.2. Forecasting Results for the Euro area

Results without pre-selection of variables to compute factors

We start with the results for the Euro area, wherein Table 2 provides information on the fraction of the variance of the panel explained by the factors. For the Euro area for the whole sample and without any pre-selection of variables, seven or eight factors are needed to explain 50% of the variance of the sample.

Figure 6 shows the adjusted R^2 for the three variables of interest, recursively computed for a sample that starts in 1991 and ends in 1997-2005. For HICP inflation, the adjusted R^2 is approximately 40% for the first two factors, rising to 50% in some periods with four factors, but with no systematic gains evident from the use of more than four factors. The corresponding panel in Figure 6 for manufacturing output growth shows that even the inclusion of many factors does not increase the adjusted R^2 far above 30%, and the first two factors play a minor role For the unemployment rate the adjusted R^2 is approximately 70% with the first two factors for the early samples, with insignificant roles for the other factors but it progressively decreases to about 40%. The latter feature provides additional evidence of instability.

In light of the empirical results reported so far and of the emphasis paid in the simulation exercise to the impact of structural instability in the data, it is natural to try and investigate further the existence of such instability in our Euro area dataset. This is especially so, given that the introduction of the euro and the lead up to this introduction happens through the middle part of our dataset.

Figures 7 – 9 plot the estimated coefficients from the fac_fdiarlag_bic model for HICP inflation, manufacturing output growth and unemployment rate, respectively. The coefficients are estimated recursively, while model specification is based on the full sample. Concentrating on the recursive tracks of the coefficient estimates on the first two factors, we can observe some instability of the corresponding coefficients. Especially evident and similar across the three variables are the increases in the recursive coefficients of the second factor at the period of introduction of the euro in 1999.

The kind of instability detected is similar to the case corresponding to time-varying loadings (λ) in the simulation experiments, even though here we have considered the y equation rather than the x_{it} equations, in the notation of Section 3. That the main source of instability of the factor structure may be related to time-variation of factor loadings can be seen also from Figure 10 that presents recursive eigenvalues. The eigenvalues are normalized by their sum such that they effectively measure the contribution of each factor to overall

[10] The unemployment rate was treated as I(0). The results are highly robust to treating it as I(1). These additional

variability of the panel. Instability of normalized eigenvalues can be thus interpreted as instability of contribution of corresponding factors to overall variation of variables and thus instability of factor loadings. What can be observed from upper panel of Figure 10 that reports the recursive eigenvalues for the Euro area is a distinct change in the second eigenvalue at the time of introduction of the Euro, which helps to explain the changes in the corresponding regression coefficient in the same period that we refer to above. Similar observations hold for the fourth eigenvalue. The second finding that emerges from Figure 10 is a significant degree of instability of virtually all recursive eigenvalues, which points to a significant degree of overall instability of factor loadings.

In a similar context, changes in the persistence of the factors (α) are another possible source of instability. Figure 11 provides the recursive estimates of factor persistence for the first two factors, estimated as the coefficient in AR(1) models, which however show relatively stable behaviour.

From the simulation experiments, we know that time-varying loadings can lead to deteriorations in the forecasting performance of factor models. To assess whether this is the case in this empirical application, Table 3 reports the results of a forecast comparison exercise, not only for the full sample but also for two sub-samples given by 1991:2 – 1998:12 and 1999:1 – 2005:10 respectively. For the full sample, the forecasting period is given by 1997:1 – 2005:10, for the first sub-sample 1997:1 – 1998:12 and for the second sub-sample 2003:11 – 2005:10. This implies that time-series dimensions we use for producing the first forecast are T=59 for the full sample and the first sub-sample, and T=58 for the second sub-sample (1999 – 2005).[11]

From Table 3, for forecasting HICP inflation no significant gains are evident from the use of factor models, when the factors are computed from the full dataset, and the comparison is made over the full sample. While providing the best forecasts, the best factor model for HICP inflation is given, for example, by fac_fdiar_bic but with a relative mean squared error of only 0.99. For manufacturing output growth the results are somewhat better with fbp_fdiarlag fbp_fdiar_bic,_bic, fac_fdiar_02, fac_fdiar_03 and fac_fdiar_06 providing gains of 13%, while f_pooled is also beneficial and provides a gain of 11%. These are again the

---

results are available from the authors upon request.

[11] Effective sample sizes are even smaller because of pre-sample values needed in all the models containing lags of forecast variable and factors.

best performing models. For unemployment however the best performing factor model is 21% worse than the benchmark. Other non-factor models also do not outperform the simple autoregressive model. Figure 12 plots the actual series and the one-step ahead forecasts obtained from the best non-factor model and the best factor model for all three variables.

The full sample results are somewhat in contrast with the outcomes when only the sub-sample 1991:2-1998:12 is considered. For HICP inflation, the best factor models are given by fac_fdiar_04 and fac_fdiar_05 that now outperform the benchmark by 45% and numerous other factor models in this class provide gains of between 40% to 35%. The corresponding results for manufacturing output growth also show some improvement in the relative forecast performance of factor models relative to forecasting for the whole sample period. One of the best factor models remains fbp_fdiar_bic but now provides a 25% improvement over the benchmark. The pooled factor forecast is 14% better than the benchmark. Finally, for unemployment, the best factor model (fac_fdiar_01) now is 2% worse than the benchmark. The non-intercept corrected pooled forecast is 18% worse than the benchmark

Turning to the results for the second sub-sample, we note that the good performance of factor models for inflation for the previous sub-sample are now again absent, and the results are roughly on par with those for the full sample. The best forecasting models are fac_fdiar_02 and fac_fdiar_05 but with gains of only 4%. For manufacturing output growth, the best forecasting model is given by fac_fdiar_05 with a gain of 13% which matches the full sample results. For unemployment, a set of factor models is best but with gains of only 1%.

Overall, these results are in line with the findings of the Monte Carlo experiments, since the forecasting performance of the factor models is worse over the full sample and the second sub-sample, which include the "breaking" periods of 1999 and 2002. Furthermore, again in line with the experiments, the AR benchmark is beaten by a factor model in most cases.

Finally, a small comment is due on the role of the devices to robustify the forecasts in the presence of structural change. From the results in Table 3, intercept corrections appear to work badly in all cases. Double differencing the variable of interest does not seem to produce any systematic gains.

Results with pre-selection of variables

Instability is seen to be one reason for the performance of factor models discussed above. However, in light of the rather low correlation of the variables of interest with the factors, noted in Table 2 and Figure 6, it is also of interest to consider whether factor-based forecasts would benefit from pre-selection of the variables used to compute the factors. We have therefore re-estimated the factors using a subset of the full dataset, which only includes those series selected by the Boivin and Ng (2005) criterion. They propose checking the correlation between the forecast variable and each of the variables in the panel. Only those variables whose correlation coefficient (with the forecast variable) exceeds a chosen threshold are then included in the subset of variables that are used for factor extraction. The forecast comparison based on the new set of factors is reported in Table 4, for the three variables and periods of interest. Details about chosen threshold values for correlation coefficients for different variables and resulting N dimensions are given in the notes to Tables 4 and 7.

The largest gains with respect to the figures reported in Table 3 are for manufacturing output growth, with the best factor models now producing gains of about 25 % with respect to the AR benchmark also when the forecast comparison is conducted over the whole sample. This is a reasonable finding, since the factors extracted from the full dataset had the lowest explanatory power for manufacturing output growth. For HICP inflation, the gains from pre-selection are in the region of 10 percentage points over the full sample and the second sub-sample. The results for unemployment match those without pre-selection.

With respect to the effects of structural changes, the factor models still in general perform worse over the full sample and the second sub-sample. This is particularly evident in the case of inflation, the variable that was likely most and quickly affected by the introduction of the euro.

4.3. Forecasting Results for Slovenia

Table 5 provides information on the fraction of the variance of the panel explained by the factors. For Slovenia for the whole sample and without any pre-selection of variables, four factors explain 50% of the variance of the sample, while seven or eight factors are needed to explain roughly 70% of the variance, which is better than the performance for the Euro area, partly reflecting the fact that the number of variables is larger in this case, 95 versus 57, and the frequency is quarterly rather than monthly. The lower panel of Figure 10 reveals that instability of factor loadings may be an important issue in forecasting also for Slovenia. In addition to short time series this increases the importance of having robust forecasting tools.

The MSE of the competing methods relative to the benchmark AR model, without Boivin-Ng pre-selection, are reported in Table 6, while Figure 13 graphs the actual values for each variable jointly with the best factor-based and non-factor-based forecast. Since the time span is insufficient to conduct any sub-sample analysis, we report only the results for the full sample given by 1994:1 – 2005:4.[12] The results are encouraging for the use of factor models, especially for core inflation where gains of up to 68% (fac_ic_fdi_04 and fac_ic_fdiar_04) are evident. The results are less impressive for CPI inflation, with a maximum gain of 18% (fbp_fdiar_bic and fbp_fdiarlag_bic). For GDP growth, while factor models do best, they provide gains of only up to 5%.

Table 7 provides the corresponding results with Boivin-Ng pre-selection. The best performing factor model now provides gains of 37% for CPI inflation, up to 14% for GDP growth and 74% for core inflation. Therefore, factor models with pre-selection are shown to be efficacious here for all the variables concerned, although the absence of any sub-sample analysis makes a comparison with the Euro area results difficult. However, the good performance in this context of double differencing and intercept corrections confirms the presence of instability in the sample under analysis.

4.4 Summary

In summary, the following points emerge from the forecast comparison exercise for our two empirical examples.

(a) For the Euro area, the detected instability is similar to the case corresponding to time-varying loadings (λ) in the simulation experiments. For forecasting HICP inflation no significant gains are evident from the use of factor models over the full sample. For manufacturing output, factor models provide gains of up to 13%, relative to the benchmark, while for unemployment the benchmark outperforms all competing models.

(b) When only the sub-sample 1991:2-1998:12 is considered, to account for the instability noted in (a), the best factor model for HICP inflation outperforms the benchmark by 45%. For manufacturing output growth, the best factor model provides a 25% improvement over the benchmark, while for unemployment, the benchmark remains the best performer.

For the second sub-sample, 1999:1 – 2005:10, more subject to changes, the results are roughly on par with those for the full sample.

(c) With variable pre-selection, according to Boivin and Ng (2006), the largest gains are for manufacturing growth, with the best factor models out-performing the benchmark by 25%, both for the full sample and the sub-samples. The performance of factor models in forecasting HICP inflation also improves for the full sample and the second sub-sample.

(d) In the case of Slovenia, where instability is expected to be even more diffuse, the results for factor models are most impressive for core inflation, and variable pre-selection improves the forecasting performance of factor models for all three variables – namely core inflation, GDP growth and CPI inflation.

5. Conclusions

In this paper we have evaluated the forecasting performance of diffusion index-based methods in short samples with structural change. Typically, factor forecasts have been computed for large datasets of long time series of macroeconomic variables, but the case of short time series is perhaps even more interesting and relevant in practical applications. Similarly, the many and frequent changes in the economic environment suggest that it is important to assess the properties of factor forecasts in the presence of structural changes in the parameters of the underlying factor model.

We have conducted a detailed simulation study, using data generation processes selected to mimic different types of structural change, and comparing the factor forecasts with more traditional time series methods. The results indicate that the most significant effects on the forecasting performance of factor models come from time-variation in factor loadings, especially for the dimensions of panels likely to be encountered by us in practice. However, in the majority of cases, factor forecasts are more accurate than standard time series forecasts, except when the sample size is very small and many factors are significant in the forecasting equation.

[12] Overall , the sample contains 48 time series observations. Because the forecasting period is 2001:3-2005:4 (see Table 6), the first model that is used to produce the first forecast for 2001:3 uses only 26 observations (minus pre-sample values to account for lags in the forecasting models).

In order to provide empirical content to our simulation analysis, we have conducted a forecast evaluation exercise for the Euro area and Slovenia. In both cases large datasets of macroeconomic time series are available, but officially only for rather short samples, likely characterized by structural changes related to the introduction of the euro in the case of the Euro area, and to accession to the European Union for Slovenia. A detailed recursive analysis showed clear evidence of instability, particularly for the Euro area forecasts and factors. As expected, most changes occurred in the time period contiguous to the introduction of the euro. In general, the factor forecasts compared well with respect to the AR competitors, while other common tools to robustify the forecasts in the presence of structural changes, such as intercept corrections and double differences, were not so useful in our context.

Overall, our results provide yet another warning on the deleterious effects of parameter changes for forecasting, but also a positive indication on the performance of factor forecasts with respect to standard time series methods in small panels with structural change.

References

Angelini, E., J. Henry and M. Marcellino (2006). Interpolation and backdating with a large information set. *Journal of Economic Dynamics and Control 30*, 2693-2724.

Artis, M. J. and M. Marcellino (2001). Fiscal forecasting: the track record of IMF, OECD and EC. *Econometrics Journal 4*, 20-36.

Artis, M.J., A. Banerjee, and M. Marcellino (2005). Factor forecasts for the UK. *Journal of Forecasting 24*, 279-298.

Bai, J. and S. Ng (2006). Confidence intervals for diffusion index forecasts with a large number of predictors. *Econometrica 74*, 1133-1150.

Banerjee, A., M. Marcellino and I. Masten (2005). Leading indicators for Euro area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics 67*, 785-813.

Banerjee, A., M. Marcellino and I. Masten (2006). Forecasting macroeconomic variables for the accession countries. In Artis, M.J. , A. Banerjee and M. Marcellino (Eds.), *The Central and Eastern European Countries and the European Union* (pp. 108-34), Cambridge: Cambridge University Press.

Boivin, Jean & Ng, Serena, 2006. "Are more data always better for factor analysis?," Journal of Econometrics, Elsevier, vol. 127(1), pp. 169-194

Breitung, J. and S. Eickmeier (2005). *Dynamic factor models*. Discussion Paper Series 1: Economic Studies No 38/2005, Deutsche Bundesbank.

Bruneau, C., O. de Bandt, A. Flageollet, and E. Michaux (2006). Forecasting Inflation using Economic Indicators: the Case of France. *Journal of Forecasting,* forthcoming.

Clements, M. and D.F. Hendry (1999). *Forecasting Non-Stationary Economic Time Series*. Cambridge, Massachusetts: MIT Press.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000). The generalised factor model: identification and estimation. *The Review of Economic and Statistics 82*, 540-554.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005). The Generalized Dynamic Factor Model one-sided estimation and forecasting. *Journal of the American Statistical Association 100*, 830-840.

Kapetanios G. and M. Marcellino (2006). *A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions*. CEPR Working Paper No 5620.

Marcellino, M. (2006). *A simple benchmark for forecasts of growth and inflation*. Unpublished manuscript.

Marcellino, M., J.H. Stock and M.W. Watson (2003). Macroeconomic forecasting in the Euro area: country specific versus euro wide information. *European Economic Review 47*, 1-18.

Marcellino, M., J.H. Stock and M.W. Watson (2006). A Comparison of Direct and Iterated AR Methods for Forecasting Macroeconomic Series h-Steps Ahead. *Journal of Econometrics 135*, 499-526.

Matheson, T. (2005). *Factor model forecasts for New Zealand.* Discussion Paper 2005/01, Reserve Bank of New Zealand.

Schumacher, C. (2006). Forecasting German GDP using alternative factor models based on large datasets. *Journal of Forecasting*, forthcoming.

Stock, J.H. and M.W. Watson (1998). *Diffusion indexes.* NBER Working Paper No. 6702.

Stock, J.H. and M.W. Watson (1999). Forecasting inflation. *Journal of Monetary* Economics *44*, 293-335.

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

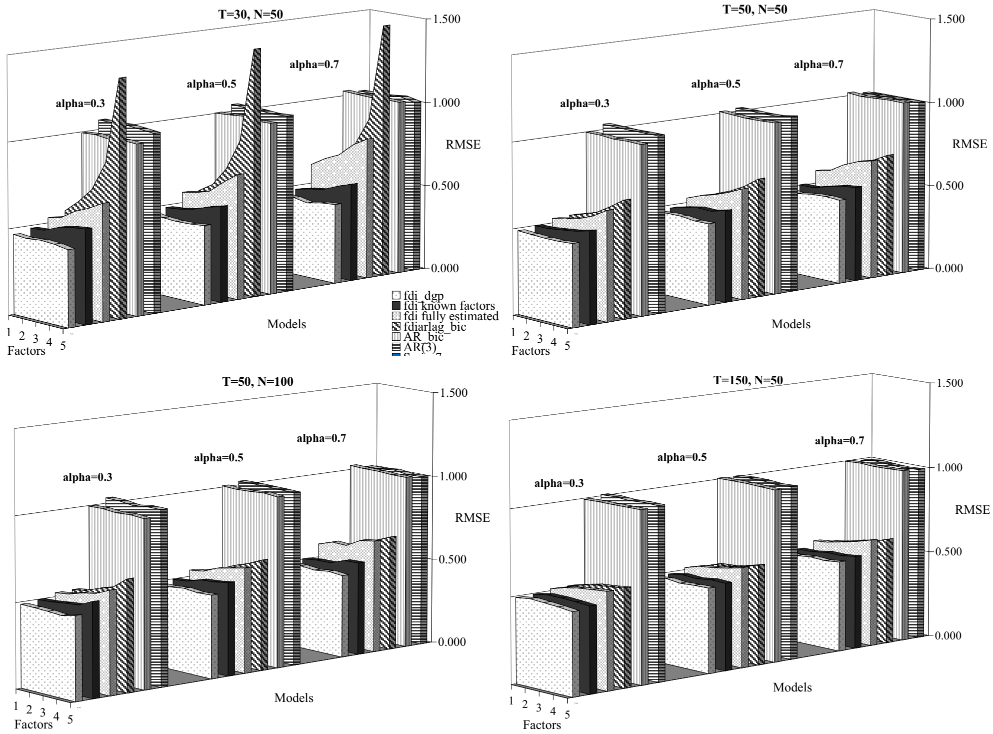
Stock, J.H. and M.W. Watson (2002b). Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics 20*, 147-162.

Stock, J.H. and M.W. Watson (2006). Forecasting with many predictors. In Elliott, G., C.W.J. Granger and A. Timmermann (Eds.) *Handbook of Economic Forecasting*, Amsterdam: North Holland.

West, K.D. (1996). Asymptotic inference about predictive ability. *Econometrica 64*, 1067-1084.

Monte Carlo Results

Figure 1: Basic DGP, *h* = 1

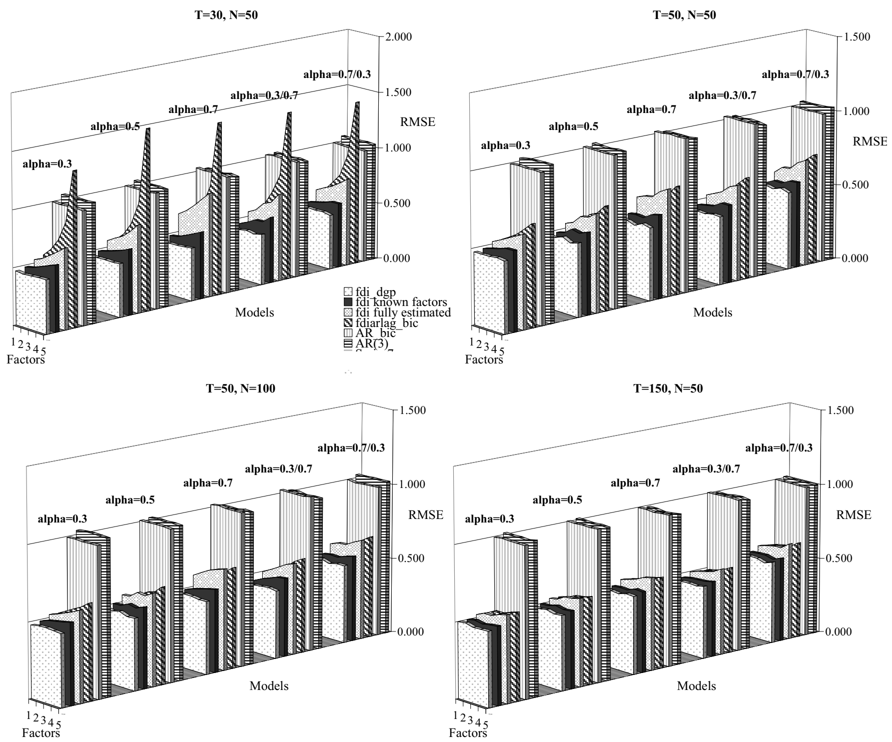


Note: The data are generated from the factor model (2) – (7), uncorrelated errors, without time-varying loadings. The depicted bar graphs for different number of factors (denoted f = 1,2,…5) are averages over 10, 000 Monte Carlo replications of the MSE for each model relative to an AR(1) benchmark, for h = 1 step-ahead forecasts..
The models under comparison are:

fdi_dgp:	known model, known parameters, known factors
fdi known factors:	known model, unknown parameters, known factors
fdi fully estimated:	known model, unknown parameters, unknown factors
fdiarlag_bic:	unknown model, parameters and factors; model selection by BIC
AR_bic	AR model, BIC lag selection
AR(3):	AR model, 3 lags

Time-varying alpha

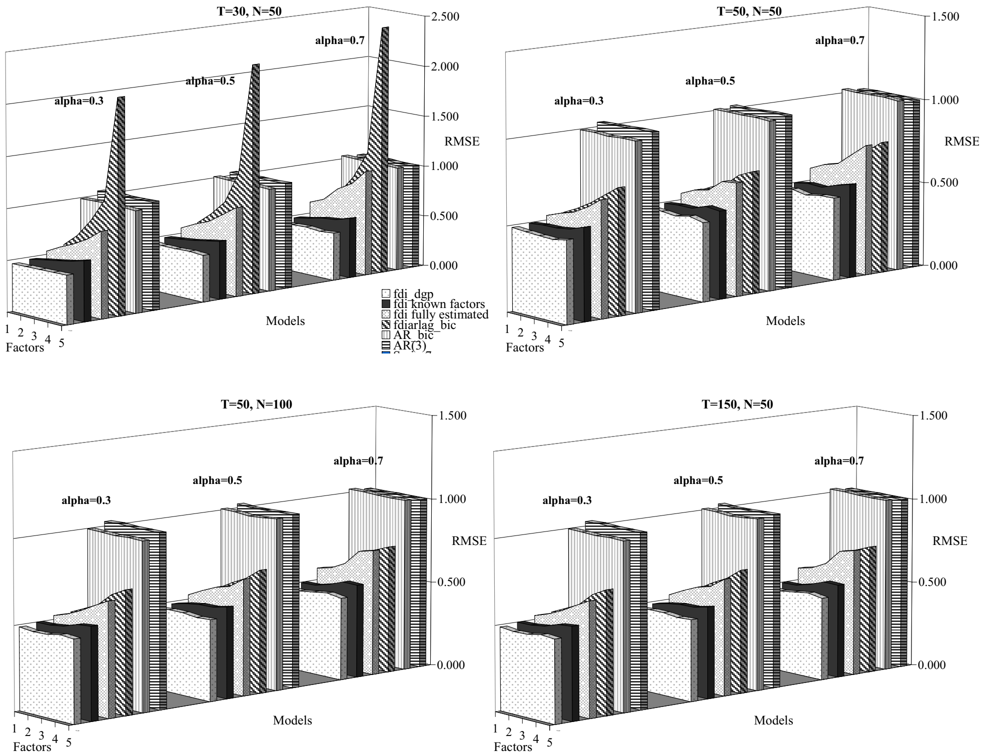
Figure 2: Continuous time variation and discrete breaks in factors persistence, *h* = 1



Note: See notes to Figure 1.

Time-varying lambda

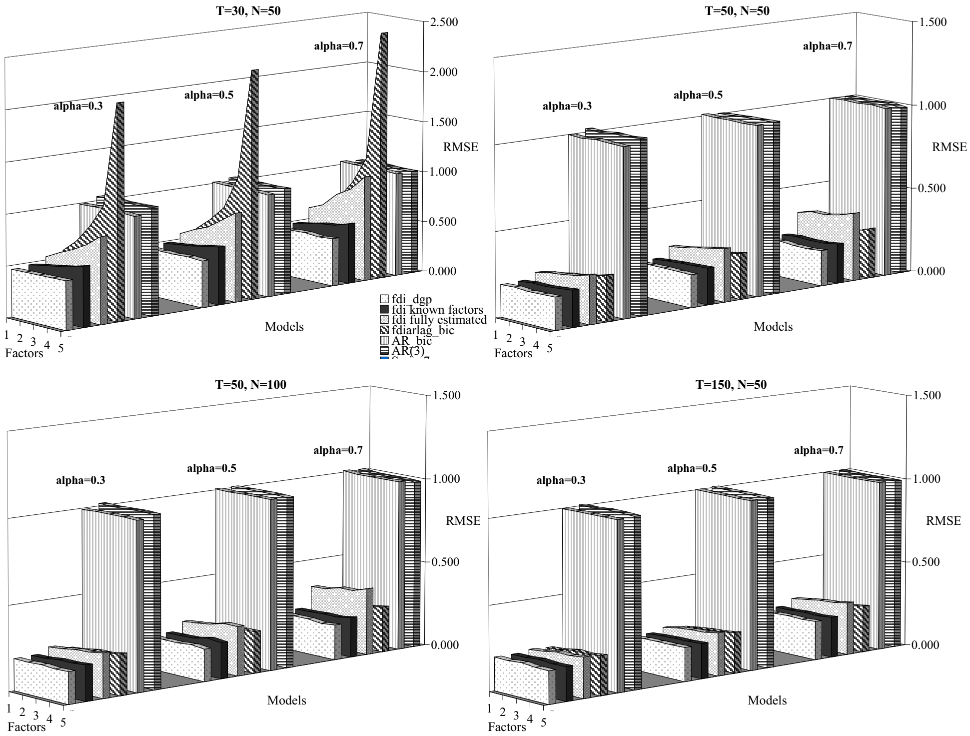
Figure 3: Time variation in factor loadings, *h* = 1



Note: See notes to Figure 1.

Double variance of factors

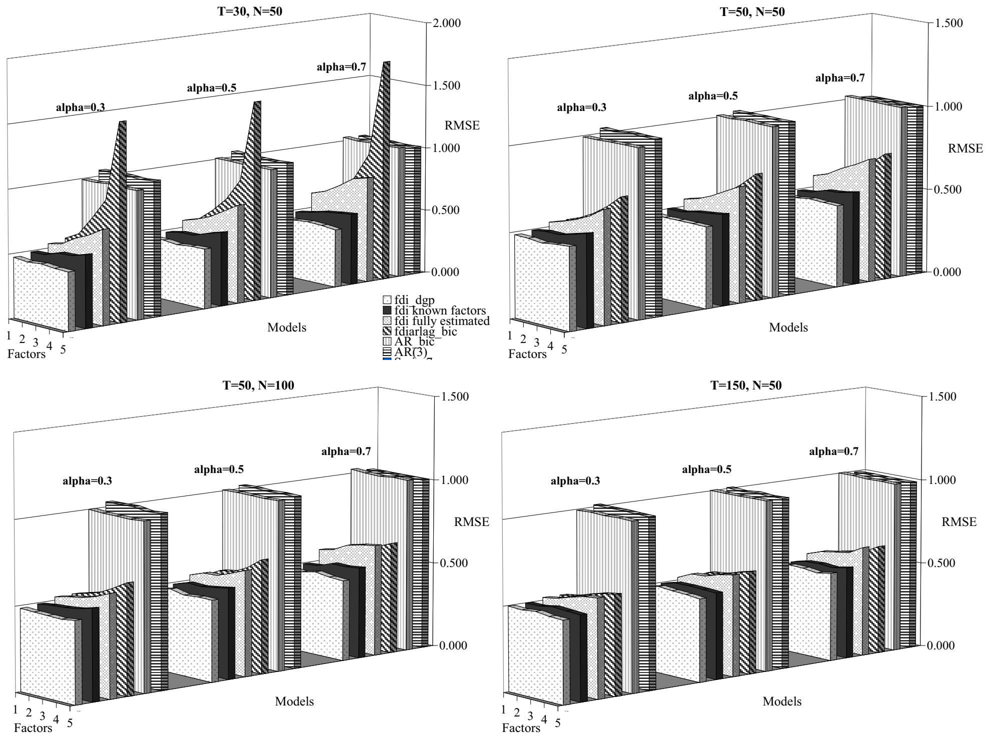
Figure 4: Double variance of factors, *h* = 1



Note: See notes to Figure 1.

Cross and time correlation of idiosyncratic components

Figure 5: Cross and time correlation of idiosyncratic components, *h* = 1



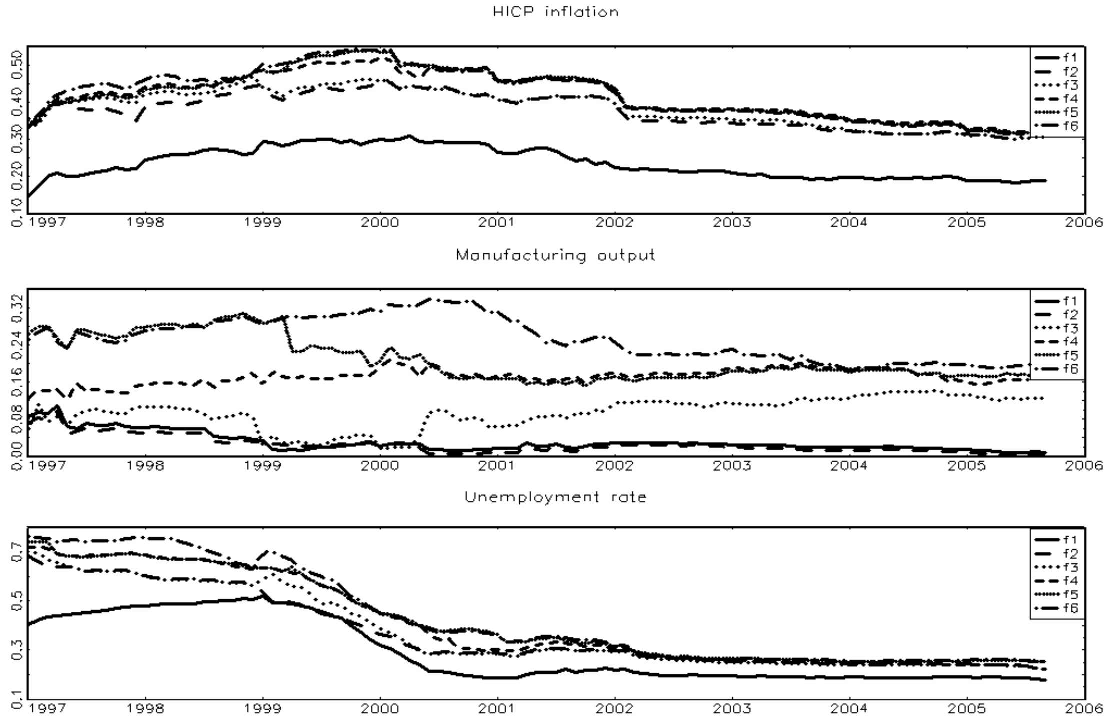
Note: See notes to Figure 1.

Empirical results

Table 2: Fraction of variance of the panel explained by the factors for the Euro area, sample 1991 - 2005

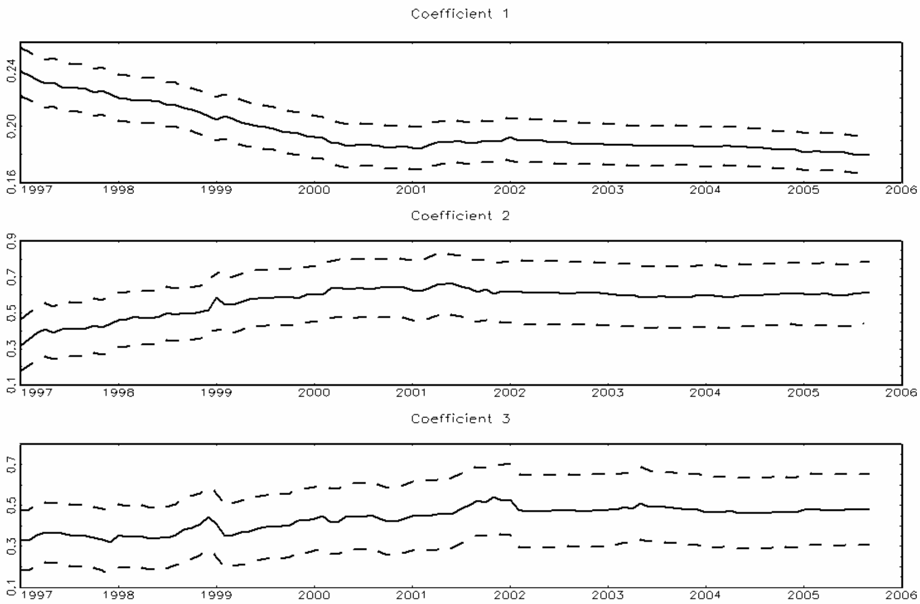
Factor	Marginal Trace R-squared	Cumulative Trace R-squared
1	0.14	0.14
2	0.09	0.23
3	0.07	0.30
4	0.06	0.36
5	0.05	0.41
6	0.04	0.45
7	0.04	0.49
8	0.04	0.52
9	0.03	0.56
10	0.03	0.59
11	0.03	0.61
12	0.03	0.64
N	57	

Figure 6: Recursive adjusted R^2, h = 1, Euro area



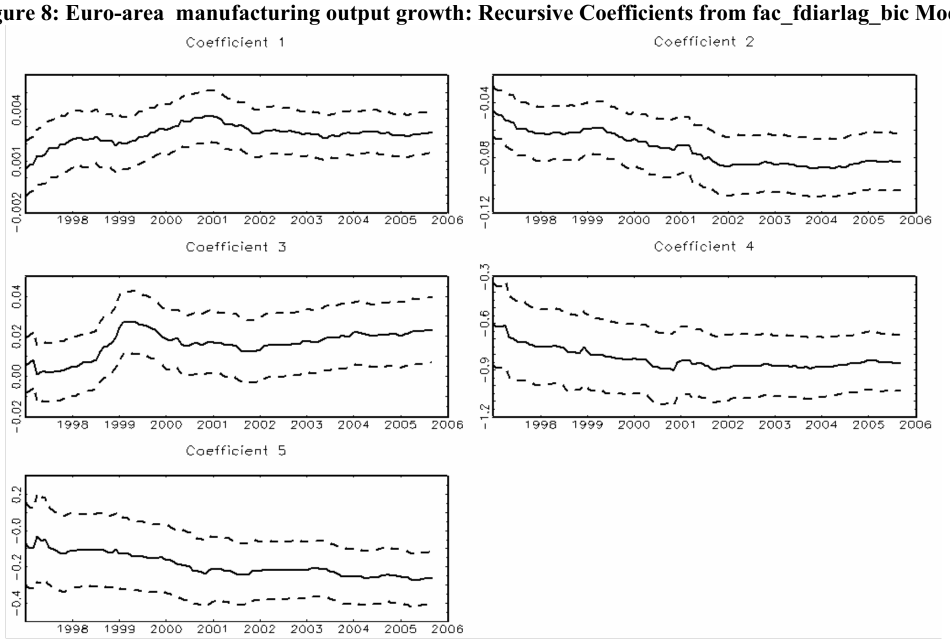
Note: Each variable is regressed on factors lagged h-period. The regressions consecutively include 1 up to 6 factors, consecutively denoted by f1, f2,…, f6..

Figure 7: Euro-area inflation: Recursive Coefficients from fac_fdiarlag_bic Model



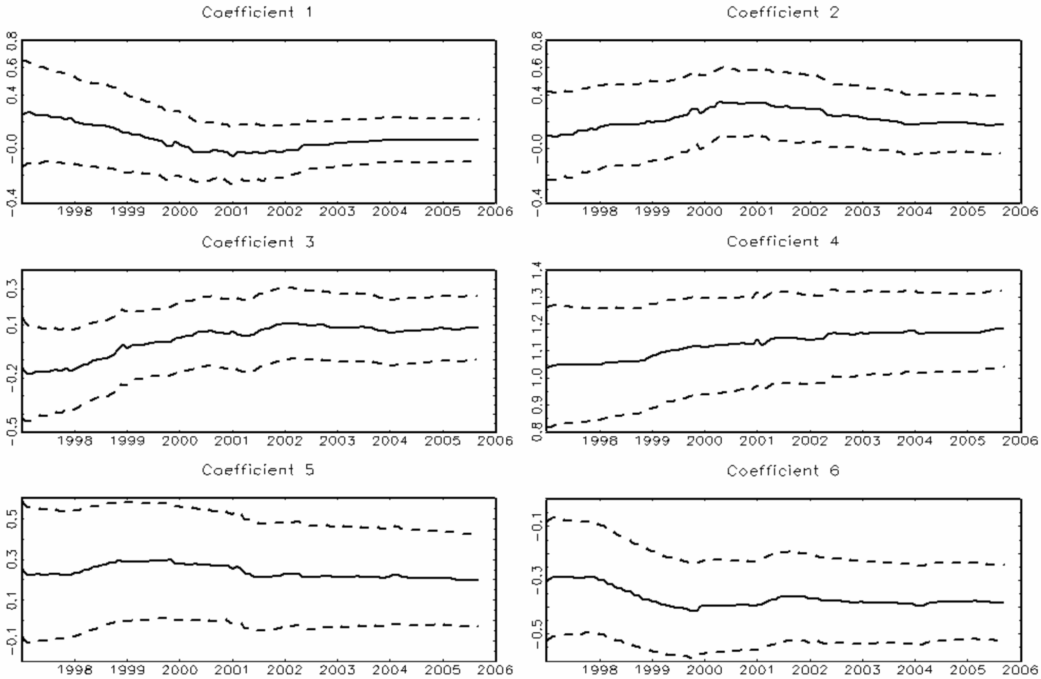
Note: The model estimated is fdiarlag_bic over the whole sample (1991-2005), which resulted in BIC eliminating all endogenous lags and keeping the first two factors lagged once (plus a constant). The first coefficient is the constant, the panel labelled Coefficient 2 is for the coefficient estimate on the first factor and the panel labeled Coefficient 3 is the coefficient estimate on the second factor. The dashed lines indicate the 95% confidence intervals.

Figure 8: Euro-area manufacturing output growth: Recursive Coefficients from fac_fdiarlag_bic Model



Note: The model estimated is fdiarlag_bic over the whole sample (1991-2005), which resulted in BIC retaining two endogenous lags and the first two factors lagged once (plus a constant). The first coefficient is for the constant, the panels labeled Coefficient 2 and Coefficient 3 refer to the first and the second factor respectively. Labels Coefficient 4 and Coefficient 5 are for the coefficients on two lags of industrial output. The dashed lines indicate the 95% confidence intervals.

Figure 9: Euro-area unemployment rate: Recursive Coefficients from fac_fdiarlag_bic Model



Note: The model estimated is fdiarlag_bic over the whole sample (1991-2005), which resulted in BIC retaining three endogenous lags (plus a constant). The first two factors lagged once were nevertheless retained in the model for comparability. The first coefficient is for the constant, the panels labeled Coefficient 2 and Coefficient 3 refer to the first and the second factor respectively. Labels Coefficient 4 - 6 are for the coefficients on three lags of unemployment rate. The dashed lines indicate the 95% confidence intervals.

Figure 10: Recursive normalized eigenvalues

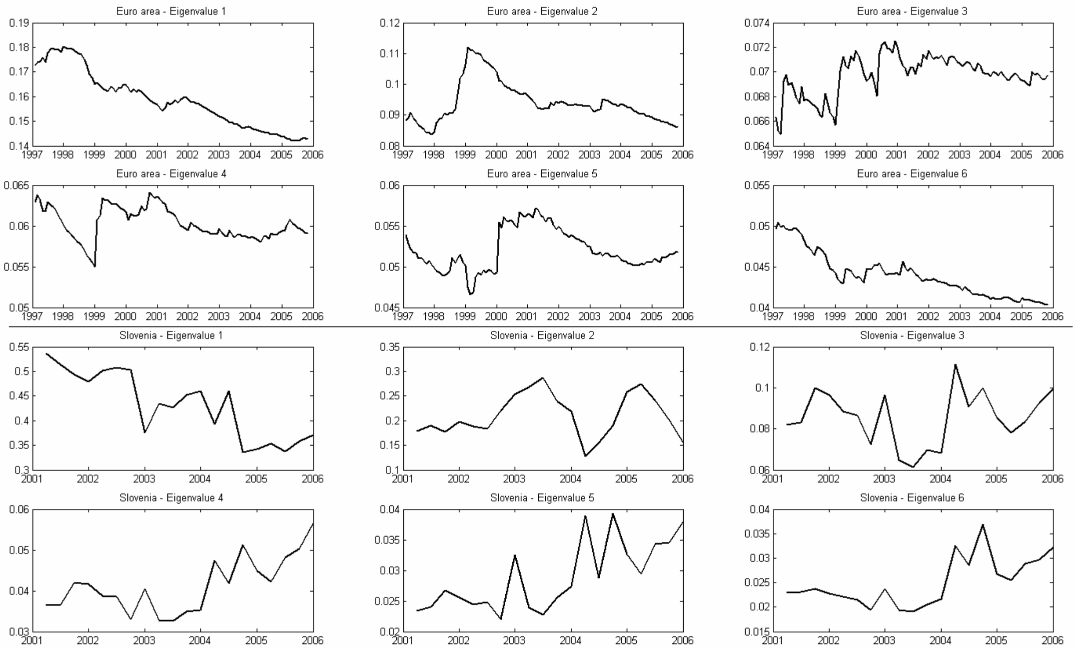
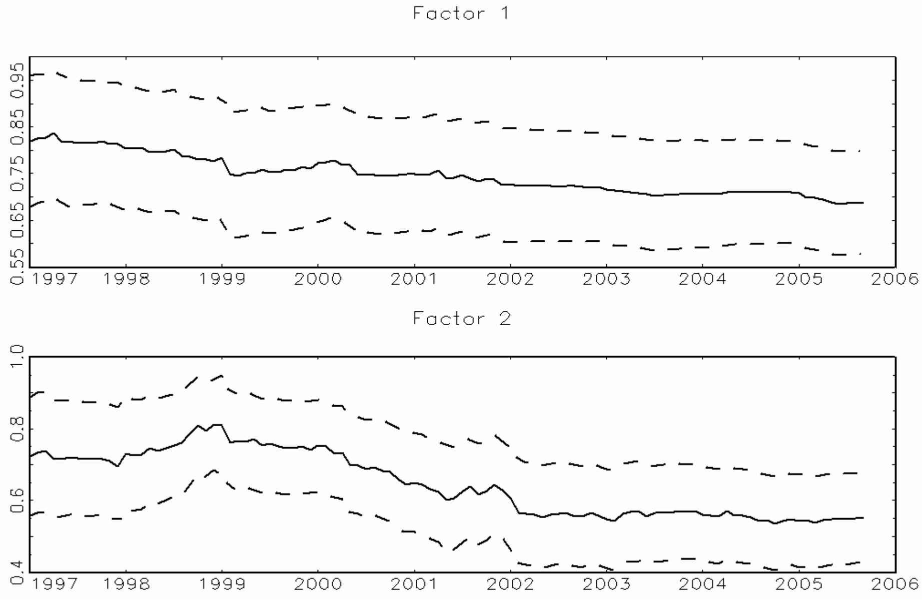


Figure 11: Recursive persistence of factors, Euro area, sample 1991 – 2005



Note: Dashed lines indicate confidence intervals.

Table 3: Effect of Euro introduction: comparison of forecast performance of factor models across sub-periods

	Inflation						Manufacturing output growth					
Forecast method	91 - 05		91 - 98		99 - 05		91 - 05		91 - 98		99 - 05	
ar_bic	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
ar_bic_i2	1.01	(0.05)	0.79	(0.14)	1.27	(0.33)	1.13	(0.09)	1.09	(0.16)	1.38	(0.35)
ar_bic_ic	2.10	(0.56)	1.85	(0.85)	2.54	(1.33)	2.28	(0.65)	2.58	(1.97)	2.61	(1.25)
varf	1.00	(0.06)	0.85	(0.12)	1.55	(0.65)	1.02	(0.07)	0.96	(0.09)	1.16	(0.17)
varfic	2.18	(0.60)	1.62	(0.58)	3.78	(3.37)	2.17	(0.58)	2.40	(1.72)	2.83	(1.96)
fac__fdiarlag_bic	0.99	(0.11)	0.63	(0.18)	1.01	(0.17)	0.89	(0.08)	0.83	(0.15)	1.00	(0.18)
fac__fdiar_bic	0.99	(0.11)	0.63	(0.18)	1.01	(0.17)	0.89	(0.08)	0.83	(0.15)	0.92	(0.17)
fbp__fdiarlag_bic	1.01	(0.12)	0.65	(0.17)	1.03	(0.18)	0.87	(0.08)	0.75	(0.15)	0.95	(0.17)
fbp__fdiar_bic	1.01	(0.12)	0.65	(0.17)	1.01	(0.17)	0.87	(0.08)	0.75	(0.15)	0.92	(0.17)
fac__fdiar_01	1.00	(0.06)	1.02	(0.14)	1.01	(0.17)	0.90	(0.08)	0.80	(0.16)	0.92	(0.17)
fac__fdiar_02	1.02	(0.10)	0.64	(0.18)	0.96	(0.18)	0.87	(0.08)	0.80	(0.13)	0.91	(0.17)
fac__fdiar_03	0.99	(0.09)	0.58	(0.18)	1.00	(0.19)	0.87	(0.08)	0.85	(0.11)	0.94	(0.18)
fac__fdiar_04	0.99	(0.12)	0.55	(0.17)	0.97	(0.18)	0.88	(0.08)	0.84	(0.12)	0.87	(0.19)
fac__fdiar_05	0.99	(0.12)	0.55	(0.17)	0.96	(0.17)	0.89	(0.07)	0.84	(0.11)	0.87	(0.18)
fac__fdiar_06	1.00	(0.12)	0.61	(0.18)	1.04	(0.17)	0.87	(0.07)	0.86	(0.11)	0.88	(0.19)
fac_ic_fdiarlag_bic	2.38	(0.86)	1.60	(0.62)	2.63	(1.32)	1.99	(0.42)	2.07	(1.06)	2.08	(0.78)
fac_ic_fdiar_bic	2.38	(0.86)	1.60	(0.62)	2.63	(1.32)	1.99	(0.42)	2.07	(1.06)	2.00	(0.70)
fbp_ic_fdiarlag_bic	2.46	(0.88)	1.59	(0.61)	2.71	(1.39)	1.93	(0.40)	1.88	(0.90)	1.93	(0.70)
fbp_ic_fdiar_bic	2.46	(0.88)	1.59	(0.61)	2.66	(1.34)	1.93	(0.40)	1.88	(0.90)	2.01	(0.71)
fac_ic_fdiar_01	2.20	(0.64)	2.05	(1.17)	2.63	(1.32)	2.02	(0.42)	1.99	(0.95)	2.00	(0.70)
fac_ic_fdiar_02	2.33	(0.80)	1.64	(0.66)	2.52	(1.25)	1.98	(0.43)	2.09	(1.09)	2.02	(0.72)
fac_ic_fdiar_03	2.27	(0.72)	1.40	(0.48)	2.71	(1.43)	1.98	(0.49)	2.22	(1.47)	2.16	(0.81)
fac_ic_fdiar_04	2.33	(0.85)	1.25	(0.40)	2.58	(1.28)	1.98	(0.48)	2.17	(1.39)	1.89	(0.66)
fac_ic_fdiar_05	2.45	(0.93)	1.21	(0.40)	2.74	(1.36)	2.04	(0.51)	2.18	(1.43)	1.90	(0.68)
fac_ic_fdiar_06	2.48	(0.93)	1.30	(0.47)	3.08	(1.96)	2.02	(0.50)	2.23	(1.51)	2.41	(1.19)
f_pooled	1.00	(0.09)	0.69	(0.17)	1.14	(0.22)	0.89	(0.07)	0.86	(0.10)	0.92	(0.13)
f_ic_pooled	1.10	(0.14)	0.62	(0.18)	1.30	(0.31)	0.98	(0.07)	1.01	(0.10)	0.96	(0.16)
RMSE for AR model	0.097		0.095		0.075		0.010		0.011		0.009	

Table 3: continued

	Unemployment					
Forecast method	**91 - 05**		**91 - 98**		**99 - 05**	
ar_bic	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
ar_bic_i2	1.00	(0.08)	1.25	(0.19)	1.11	(0.17)
ar_bic_ic	1.64	(0.37)	2.08	(0.56)	2.03	(1.27)
varf	1.05	(0.11)	1.44	(0.29)	1.31	(0.29)
varfic	1.82	(0.35)	2.50	(0.91)	2.44	(1.46)
fac__fdiarlag_bic	1.26	(0.16)	1.11	(0.17)	0.99	(0.01)
fac__fdiar_bic	1.26	(0.16)	1.11	(0.17)	0.99	(0.01)
fbp__fdiarlag_bic	1.24	(0.15)	1.11	(0.17)	0.99	(0.01)
fbp__fdiar_bic	1.24	(0.15)	1.11	(0.17)	0.99	(0.01)
fac__fdiar_01	1.21	(0.16)	1.02	(0.12)	0.97	(0.04)
fac__fdiar_02	1.33	(0.19)	1.61	(0.38)	0.99	(0.04)
fac__fdiar_03	1.34	(0.19)	1.73	(0.38)	1.02	(0.10)
fac__fdiar_04	1.33	(0.24)	1.88	(0.65)	1.03	(0.08)
fac__fdiar_05	1.28	(0.22)	1.65	(0.44)	1.04	(0.08)
fac__fdiar_06	1.31	(0.23)	1.73	(0.50)	1.10	(0.10)
fac_ic_fdiarlag_bic	2.56	(0.95)	1.79	(0.55)	2.02	(1.26)
fac_ic_fdiar_bic	2.56	(0.95)	1.79	(0.55)	2.02	(1.26)
fbp_ic_fdiarlag_bic	2.51	(0.90)	1.79	(0.55)	2.02	(1.26)
fbp_ic_fdiar_bic	2.51	(0.90)	1.79	(0.55)	2.02	(1.26)
fac_ic_fdiar_01	2.48	(0.90)	1.74	(0.47)	2.02	(1.24)
fac_ic_fdiar_02	2.45	(0.90)	1.73	(0.43)	2.05	(1.24)
fac_ic_fdiar_03	2.46	(0.90)	1.97	(0.66)	1.90	(0.92)
fac_ic_fdiar_04	2.30	(1.01)	2.08	(0.82)	2.01	(1.05)
fac_ic_fdiar_05	2.18	(0.97)	1.57	(0.51)	1.98	(1.04)
fac_ic_fdiar_06	2.29	(1.03)	1.87	(0.63)	2.09	(1.16)
f_pooled	1.15	(0.13)	1.18	(0.16)	1.00	(0.05)
f_ic_pooled	1.39	(0.29)	1.58	(0.46)	0.99	(0.09)
RMSE for AR model	0.065		0.060		0.056	

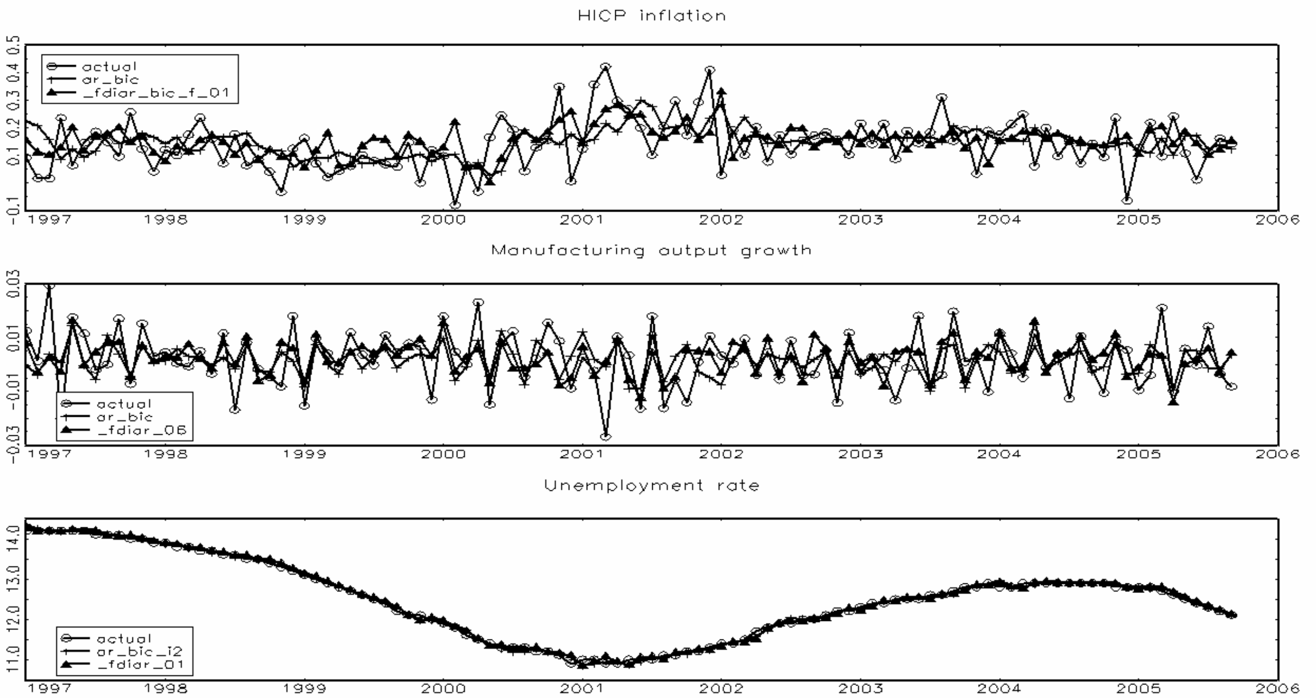
Forecasting periods for each sub-sample are: 1997:1 – 1998:12 for 1991-19 98, 2003:11 – 2005:10 for 1999 – 2005 and 1997:1 – 2005:10 for the full sample 1991 – 2005.

Notes: One-step-ahead forecasts. For each variable and sub-period columns report the MSE relative to the benchmark AR model, with West (1996) standard error in parentheses. We also report the root MSE for the AR benchmark in the last line of the table.

The forecasts in the rows of tables are (see section 2.1 for details):

ar_bic	AR model (BIC selection), benchmark
ar_bic_i2	AR model (BIC selection) for second-differenced variable
ar_bic_ic	AR model (BIC selection) with intercept correction
varf	VAR model
varfic	VAR model with intercept correction
fac__fdiarlag_bic	Factors from unbalanced panel (BIC selection), their lags, and AR terms
fac__fdiar_bic	Factors from unbalanced panel (BIC selection), and AR terms
fac__fdi_bic	Factors from unbalanced panel (BIC selection)
fbp__fdiarlag_bic	Factors from balanced panel (BIC selection), their lags, and AR terms
fbp__fdiar_bic	Factors from balanced panel (BIC selection), and AR terms
fbp__fdi_bic	Factors from balanced panel (BIC selection)
fac__fdiar_n	n factors from unbalanced panel, n = 1,…,6 and AR terms
fac__fdi_n	n factors from unbalanced panel, n = 1,…,6
fac_ic_fdiarlag_bic	As factor models above, but with intercept correction
fac_ic_fdiar_bic	
fac_ic_fdi_bic	
fbp_ic_fdiarlag_bic	
fbp_ic_fdiar_bic	
fbp_ic_fdi_bic	
fac_ic_fdiar_n	
fac_ic_fdi_n	
f_pooled	Average of factor forecasts without intercept correction
f_ic_pooled	Average of factor forecasts with intercept correction

Figure 12: Forecasting macroeconomic variables for the Euro area, sample 1991 – 2005



Note: Each figure plots the actual series and the one-step ahead forecasts obtained from the best non-factor model and the best factor model. (See note to Table 3 for definitions of forecasting methods.)

Table 4: Effect of Euro introduction: comparison of forecast performance of factor models across sub-periods with Boivin & Ng pre-selection of variables

	Inflation						Manufacturing output growth					
Forecast method	91 - 05		91 - 98		99 - 05		91 - 05		91 - 98		99 - 05	
ar_bic	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
ar_bic_i2	1.01	(0.05)	0.79	(0.14)	1.27	(0.33)	1.13	(0.09)	1.09	(0.16)	1.38	(0.35)
ar_bic_ic	2.10	(0.56)	1.85	(0.85)	2.54	(1.33)	2.28	(0.65)	2.58	(1.97)	2.61	(1.25)
fac__fdiarlag_bic	0.89	(0.09)	0.64	(0.15)	0.86	(0.19)	0.76	(0.08)	0.70	(0.16)	0.85	(0.22)
fac__fdiar_bic	0.89	(0.09)	0.64	(0.15)	0.87	(0.18)	0.76	(0.08)	0.70	(0.16)	0.88	(0.22)
fbp__fdiarlag_bic	0.92	(0.08)	0.66	(0.14)	0.86	(0.19)	0.74	(0.08)	0.68	(0.16)	0.75	(0.20)
fbp__fdiar_bic	0.92	(0.08)	0.66	(0.14)	0.87	(0.18)	0.74	(0.08)	0.68	(0.16)	0.86	(0.24)
fac__fdiar_01	0.89	(0.09)	0.64	(0.15)	0.87	(0.18)	0.76	(0.08)	0.70	(0.16)	0.88	(0.23)
fac__fdiar_02	0.89	(0.09)	0.64	(0.15)	0.95	(0.20)	0.78	(0.08)	0.74	(0.14)	0.86	(0.22)
fac_ic_fdiarlag_bic	2.15	(0.64)	1.43	(0.51)	2.05	(0.89)	1.68	(0.29)	1.81	(0.75)	1.76	(0.64)
fac_ic_fdiar_bic	2.15	(0.64)	1.43	(0.51)	2.16	(1.00)	1.68	(0.29)	1.81	(0.75)	1.88	(0.71)
fbp_ic_fdiarlag_bic	2.18	(0.64)	1.47	(0.53)	2.05	(0.89)	1.63	(0.27)	1.80	(0.78)	1.54	(0.40)
fbp_ic_fdiar_bic	2.18	(0.64)	1.47	(0.53)	2.16	(1.00)	1.63	(0.27)	1.80	(0.78)	1.85	(0.72)
fac_ic_fdiar_01	2.15	(0.64)	1.43	(0.51)	2.16	(1.00)	1.69	(0.29)	1.81	(0.75)	1.85	(0.71)
fac_ic_fdiar_02	2.16	(0.67)	1.37	(0.46)	2.30	(1.23)	1.71	(0.31)	1.93	(0.90)	1.82	(0.69)
f_pooled	0.96	(0.08)	0.72	(0.15)	1.00	(0.17)	0.85	(0.06)	0.88	(0.09)	0.95	(0.14)
f_ic_pooled	1.61	(0.33)	1.06	(0.30)	1.59	(0.51)	1.25	(0.14)	1.36	(0.32)	1.30	(0.32)
RMSE for AR model	0.097		0.095		0.075		0.010		0.011		0.009	

Table 4: continued

	Unemployment					
Forecast method	91 - 05		91 - 98		99 - 05	
ar_bic	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
ar_bic_i2	1.00	(0.08)	1.25	(0.19)	1.11	(0.17)
ar_bic_ic	1.64	(0.37)	2.08	(0.56)	2.03	(1.27)
fac__fdiarlag_bic	1.25	(0.18)	1.11	(0.17)	0.99	(0.01)
fac__fdiar_bic	1.25	(0.18)	1.11	(0.17)	0.99	(0.01)
fbp__fdiarlag_bic	1.25	(0.19)	1.11	(0.17)	0.99	(0.01)
fbp__fdiar_bic	1.25	(0.19)	1.11	(0.17)	0.99	(0.01)
fac__fdiar_01	1.19	(0.18)	0.97	(0.11)	0.99	(0.02)
fac__fdiar_02	1.20	(0.18)	1.11	(0.16)	0.99	(0.03)
fac_ic_fdiarlag_bic	2.41	(0.94)	1.79	(0.55)	2.02	(1.26)
fac_ic_fdiar_bic	2.41	(0.94)	1.79	(0.55)	2.02	(1.26)
fbp_ic_fdiarlag_bic	2.38	(0.92)	1.79	(0.55)	2.02	(1.26)
fbp_ic_fdiar_bic	2.38	(0.92)	1.79	(0.55)	2.02	(1.26)
fac_ic_fdiar_01	2.36	(0.91)	1.67	(0.43)	2.04	(1.28)
fac_ic_fdiar_02	2.36	(0.92)	1.76	(0.52)	2.08	(1.34)
f_pooled	1.11	(0.14)	1.12	(0.14)	1.03	(0.09)
f_ic_pooled	1.80	(0.51)	1.33	(0.26)	1.53	(0.62)
RMSE for AR model	0.066		0.060		0.056	

Note: Forecasting periods for each sub-sample are the same as in Table 3. Pre-selection is done by checking the correlation (in absolute value) between the forecast variable and indicators in the panel. The limit for correlation coefficient was set to 0.25 for inflation, and 0.15 for manufacturing output growth and unemployment rate. Depending on the recursively updated samples, this criterion left around 20 series in the panels for inflation and unemployment, and between 15 and 20 for manufacturing output growth. (Detailed sizes of panels available upon request.) See also notes to Table 3.

Table 5: Fraction of variance of the panel explained by the factors (whole sample)

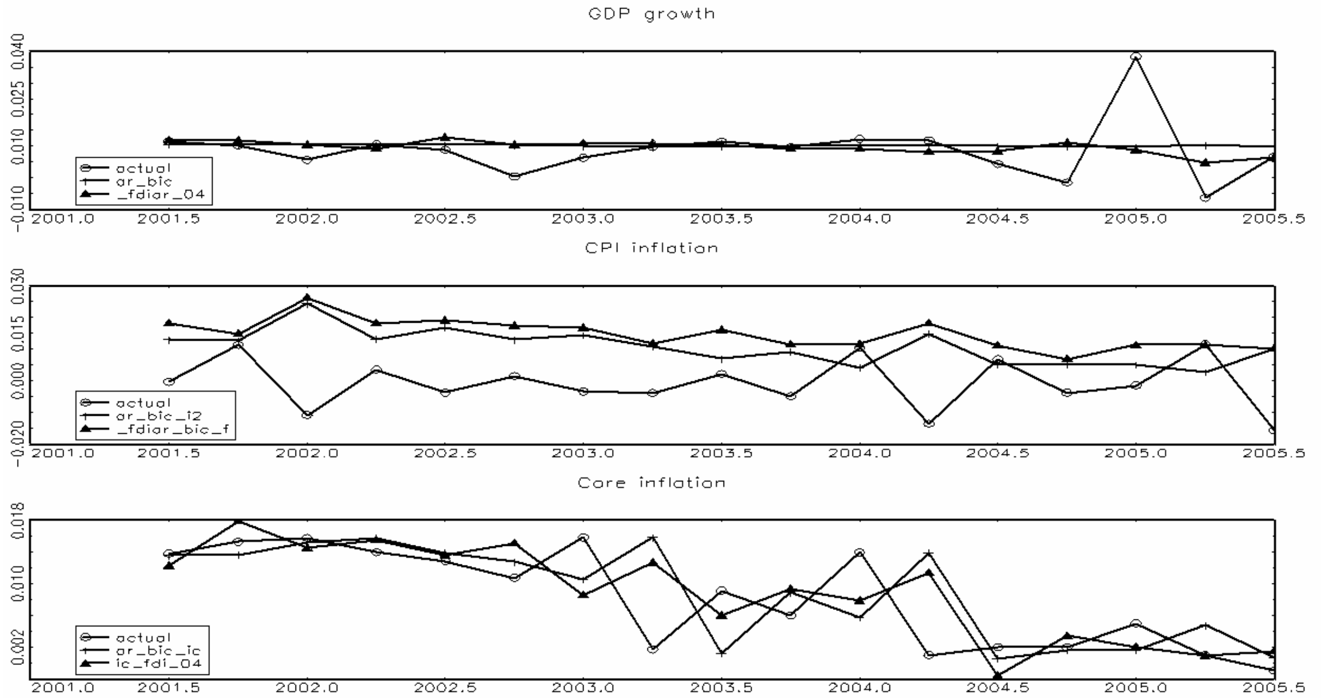
Factor	Marginal Trace R-squared	Cumulative Trace R-squared
1	0.19	0.19
2	0.14	0.33
3	0.10	0.43
4	0.08	0.51
5	0.06	0.56
6	0.05	0.61
7	0.04	0.66
8	0.04	0.70
9	0.03	0.72
10	0.03	0.75
11	0.02	0.78
12	0.02	0.80
N	95	

Table 6: Slovenia, relative MSE, *h* = 1, quarterly data, sample: 1994:1 – 2005:4, forecasting: 2001:3 – 2005:4

Forecast Method	GDP growth		CPI inflation		Core inflation	
ar_bic	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
ar_bic_i2	2.92	(4.50)	0.83	(0.27)	0.45	(0.23)
ar_bic_ic	2.91	(4.50)	1.63	(0.59)	0.45	(0.22)
fac__fdiarlag_bic	1.00	(0.00)	1.12	(0.14)	0.62	(0.19)
fac__fdiar_bic	1.00	(0.00)	1.12	(0.14)	0.93	(0.05)
fac__fdi_bic	1.00	(0.00)	1.06	(0.19)	0.93	(0.05)
fbp__fdiarlag_bic	1.00	(0.00)	0.82	(0.19)	0.72	(0.14)
fbp__fdiar_bic	1.00	(0.00)	0.82	(0.19)	0.79	(0.12)
fac__fdiar_01	1.03	(0.02)	1.22	(0.19)	0.84	(0.08)
fac__fdiar_02	1.11	(0.10)	1.27	(0.25)	0.82	(0.09)
fac__fdiar_03	0.96	(0.10)	1.10	(0.12)	0.82	(0.12)
fac__fdiar_04	0.95	(0.11)	1.14	(0.15)	0.81	(0.13)
fac__fdiar_05	0.95	(0.20)	1.15	(0.20)	0.89	(0.11)
fac__fdiar_06	0.99	(0.22)	0.95	(0.22)	0.60	(0.18)
fac__fdi_01	1.03	(0.02)	1.22	(0.19)	0.84	(0.08)
fac__fdi_02	1.11	(0.10)	1.27	(0.25)	0.82	(0.09)
fac__fdi_03	0.96	(0.10)	1.10	(0.12)	0.82	(0.12)
fac__fdi_04	0.95	(0.11)	1.14	(0.15)	0.81	(0.13)
fac__fdi_05	0.95	(0.20)	1.20	(0.22)	0.89	(0.11)
fac__fdi_06	0.99	(0.22)	0.95	(0.22)	0.60	(0.18)
fac_ic_fdiarlag_bic	2.91	(4.50)	1.65	(0.68)	0.50	(0.24)
fac_ic_fdiar_bic	2.91	(4.50)	1.65	(0.68)	0.53	(0.21)
fac_ic_fdi_bic	2.91	(4.50)	1.05	(0.37)	0.53	(0.21)
fbp_ic_fdiarlag_bic	2.91	(4.50)	1.79	(0.89)	0.72	(0.24)
fbp_ic_fdiar_bic	2.91	(4.50)	1.79	(0.89)	0.48	(0.22)
fbp_ic_fdi_bic	2.91	(4.50)	1.72	(0.84)	0.48	(0.22)
fac_ic_fdiar_01	2.97	(4.68)	1.31	(0.40)	0.50	(0.21)
fac_ic_fdiar_02	3.13	(5.31)	1.31	(0.43)	0.40	(0.22)
fac_ic_fdiar_03	2.75	(3.70)	1.46	(0.49)	0.35	(0.23)
fac_ic_fdiar_04	2.72	(3.79)	1.48	(0.50)	0.32	(0.23)
fac_ic_fdiar_05	2.52	(2.97)	1.49	(0.50)	0.50	(0.23)
fac_ic_fdiar_06	2.45	(2.77)	1.65	(0.80)	0.60	(0.23)
fac_ic_fdi_01	2.97	(4.68)	1.31	(0.40)	0.50	(0.21)
fac_ic_fdi_02	3.13	(5.31)	1.31	(0.43)	0.40	(0.22)
fac_ic_fdi_03	2.75	(3.70)	1.46	(0.49)	0.35	(0.23)
fac_ic_fdi_04	2.72	(3.79)	1.48	(0.50)	0.32	(0.23)
fac_ic_fdi_05	2.52	(2.97)	1.41	(0.48)	0.50	(0.23)
fac_ic_fdi_06	2.45	(2.77)	1.65	(0.80)	0.60	(0.23)
f_pooled	1.23	(0.33)	0.90	(0.12)	0.60	(0.16)
f_ic_pooled	1.16	(0.20)	0.86	(0.17)	0.59	(0.18)
RMSE for AR Model	0.009		0.008		0.009	

Note: See notes to Table 3.

Figure 13: Forecasting macroeconomic variables for Slovenia



Note: Each figure plots the actual series and the one-step ahead forecasts obtained from the best non-factor model and the best factor model. (See note to Table 3 for definitions of forecasting methods.)

1.000

Table 7: Boivin and Ng (2004) pre-selection of variables in the calculation of factors – Slovenia

Forecast method	GDP growth		CPI inflation		Core inflation	
ar_bic	1.00	(0.00)	1.00	(0.00)	1.00	(0.00)
ar_bic_i2	2.92	(4.50)	0.83	(0.27)	0.45	(0.23)
ar_bic_ic	2.91	(4.50)	1.63	(0.59)	0.45	(0.22)
fac__fdiarlag_bic	0.90	(0.23)	0.66	(0.17)	0.35	(0.22)
fac__fdiar_bic	0.90	(0.23)	0.66	(0.17)	0.38	(0.21)
fac__fdi_bic	0.90	(0.23)	0.63	(0.18)	0.38	(0.21)
fbp__fdiarlag_bic	1.04	(0.07)	0.66	(0.17)	0.48	(0.20)
fbp__fdiar_bic	1.04	(0.07)	0.66	(0.17)	0.55	(0.17)
fbp__fdi_bic	1.04	(0.07)	0.66	(0.17)	0.55	(0.17)
fac__fdiar_01	1.02	(0.10)	0.68	(0.17)	0.42	(0.20)
fac__fdiar_02	0.86	(0.24)	0.63	(0.18)	0.39	(0.21)
fac__fdi_01	1.02	(0.10)	0.66	(0.18)	0.42	(0.20)
fac__fdi_02	0.86	(0.24)	0.63	(0.18)	0.39	(0.21)
fac_ic_fdiarlag_bic	2.23	(2.31)	1.24	(0.38)	0.32	(0.23)
fac_ic_fdiar_bic	2.23	(2.31)	1.24	(0.38)	0.27	(0.23)
fac_ic_fdi_bic	2.23	(2.31)	1.18	(0.35)	0.27	(0.23)
fbp_ic_fdiarlag_bic	2.86	(4.16)	1.33	(0.43)	0.60	(0.22)
fbp_ic_fdiar_bic	2.86	(4.16)	1.33	(0.43)	0.52	(0.21)
fbp_ic_fdi_bic	2.86	(4.16)	1.33	(0.43)	0.52	(0.21)
fac_ic_fdiar_01	2.87	(4.01)	1.27	(0.42)	0.42	(0.22)
fac_ic_fdiar_02	2.09	(2.03)	1.13	(0.32)	0.26	(0.23)
fac_ic_fdi_01	2.87	(4.01)	1.22	(0.38)	0.42	(0.22)
fac_ic_fdi_02	2.09	(2.03)	1.13	(0.32)	0.26	(0.23)
f_pooled	1.12	(0.13)	0.67	(0.16)	0.36	(0.21)
f_ic_pooled	2.03	(1.83)	1.03	(0.25)	0.31	(0.23)
RMSE for AR model	0.009		0.008		0.009	

Note: For the correlation threshold (in absolute value) the following values were used: CPI inflation – 0.25, GDP growth – 0.20, Core inflation – 0.25. Depending on the recursively updated samples, this criterion left around 20 – 30 series in the panels. (Detailed sizes of panels available upon request.) See also notes to Table 3.